GEORGIA INSTITUTE OF TECHNOLOGY — OFFICE OF CONTRACT ADMINISTRATION

PROJECT ADMINISTRATION DATA SHEET

[X] ORIGINAL [ ] REVISION NO. ____

Project No. G-41-B06 (R5212-6A0) GTRC/~~GIT~~ XXXX DATE 8 / 22 / 85

Project Director: R. M. Wartell School/~~Lab~~ XXX Phy

Sponsor: DHHS/PHS/NIH/NIGMS

Type Agreement: Grant 5R01 GM33543-06

Award Period: From 9/1/85 To 8/31/86 (Performance) 11/31/86 (Reports)

Sponsor Amount:	This Change	Total to Date
Estimated:	$	$ 98,103
Funded:	$	$ 98,103

Cost Sharing Amount: $ 4,905 Cost Sharing No: G-41-351

Title: DNA Conformation and Protein-DNA Interaction

ADMINISTRATIVE DATA OCA Contact John Schonk x4820

1) Sponsor Technical Contact:	2) Sponsor Admin/Contractual Matters:
Helen Sunshine	Dona McNish/Barbara Spinks
National Institute of Health	National Institute of Health
NIGMS	NIGMS
Program Management	Grants Management
Bethesda, MD 20205	Bethesda, MD 20205
301/496-7125	301/496-7166

Defense Priority Rating: N/A Military Security Classification: N/A

(or) Company/Industrial Proprietary: N/A

RESTRICTIONS

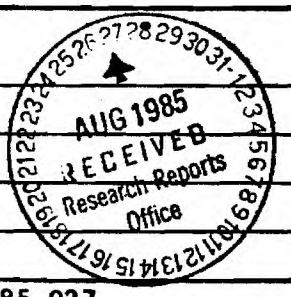
See Attached NIH Supplemental Information Sheet for Additional Requirements.

Travel: Foreign travel must have prior approval – Contact OCA in each case. Domestic travel requires sponsor approval where total will exceed greater of $500 or 125% of approved proposal budget category.

Equipment: Title vests with GIT

COMMENTS:

No funds may be expended after 8/31/86.



Continuation of G-41-B05

COPIES TO: SPONSOR'S I. D. NO. 02.108.001.85.027

Project Director	Procurement/GTRI Supply Services	GTRC
Research Administrative Network	Research Security Services	Library
Research Property Management	Reports Coordinator (OCA)	Project File
Accounting	Research Communications (2)	Other A. Jones

FORM OCA 65:285

GEORGIA INSTITUTE OF TECHNOLOGY OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION/CLOSEOUT SHEET

Date 10/28/86

Project No. G-41-B06 School/~~Lab~~ Physics

Includes Subproject No.(s) N/A

Project Director(s) R. M. Wartell GTRC /~~XHX~~

Sponsor DHHS/PHS/NIH/NIGMS

Title DNA Conformation and Protein-DNA Interaction

Effective Completion Date: 8/31/86 (Performance) 11/30/86 (Reports)

Grant/Contract Closeout Actions Remaining:

- [ ] None
- [x] Final Invoice or Final Fiscal Report
- [ ] Closing Documents
- [ ] Final Report of Inventions
- [ ] Govt. Property Inventory & Related Certificate
- [ ] Classified Material Certificate
- [ ] Other _____

Continues Project No. _____ Continued by Project No. G41-625

COPIES TO:

Project Director
Research Administrative Network
Research Property Management
Accounting
Procurement/GTRI Supply Services
Research Security Services
Reports Coordinator (OCA)
Legal Services

Library
GTRC
Research Communications (2)
Project File
Other I. Newton
A. Jones
R. Embry

FORM OCA 69.285

Volume 13 Number 20 1985 Nucleic Acids Research

The catabolite activator protein stabilizes its binding site in the *E. coli* lactose promoter

Henry DeGrazia, Saras Abhiraman and Roger M.Wartell

Schools of Applied Biology and Physics, Georgia Institute of Technology, Atlanta, GA 30332, USA

Received 22 July 1985; Accepted 30 September 1985

ABSTRACT

The effect of catabolite activator protein, CAP, on the thermal stability of DNA was examined. Site specific binding was studied with a 62 bp DNA restriction fragment containing the primary CAP site of the E. coli lactose (lac) promoter. A 144 bp DNA containing the lac promoter region and a 234 bp DNA from the pBR322 plasmid provided other DNA sites. Thermal denaturation of protein-DNA complexes was carried out in a low ionic strength solvent with 40% dimethyl sulfoxide, DMSO. In this solvent free DNA denatured below the denaturation temperature of CAP. The temperature stability of CAP for site specific binding was monitored using an acrylamide gel electrophoresis assay. Results show that both specific and non-specific CAP binding stabilize duplex DNA. Site specific binding to the 62 bp DNA produced a 13.3°C increase in the transition under conditions where non-specific binding stabilized this DNA by 2-3°C.

INTRODUCTION

The catabolite gene activator protein, CAP, in a complex with its cofactor cAMP, regulates transcription initiation of many Escherichia coli operons. This regulation is brought about by the binding of CAP-cAMP to specific DNA sites within catabolite sensitive promoters (1). The complex of CAP-cAMP and its promoter site generally enhances mRNA synthesis by RNA polymerase. The mechanism by which CAP-cAMP enhances transcription is not completely understood. DNA binding sites of CAP-cAMP are approximately 20 bp in length. Their location relative to the startpoint for transcription varies among different promoters. For the gal promoter, CAP-cAMP binds to a sequence between -25 and -50 (2). The lac promoter binding site is between -54 and -72 (3,4,5), and for the ara P_c promoter the CAP-cAMP binding site is between -78 and -107 (6). This diversity in the location of the CAP-cAMP binding sites relative to the startpoint suggests that the influence of the CAP-DNA complex on the RNA polymerase open promoter complex is variable. The ability of CAP-cAMP to act as a repressor of transcription for some promoters (7) must also be considered in deducing the mechanism(s) of CAP-cAMP's enhancement of

transcription.

Two models were initially proposed to explain the activation mechanism of CAP-cAMP. These models are not mutually exclusive. One model proposed that CAP-cAMP, when bound to DNA, interacts directly with RNA polymerase (8). A second model suggested that CAP-cAMP binding destabilizes the RNA polymerase binding site (9,10). Recent studies suggest a third model to explain CAP activation. Experiments indicate that CAP-cAMP stimulates mRNA expression of the gal and lac operons by blocking RNA polymerase from a strong binding but weak transcription site, thus allowing polymerase to bind a nearby productive initiation site (11,12,13).

The possibility that CAP-cAMP may act as a destabilizing protein in its site specific binding mode was examined by Unger et al. (14). A 301 bp fragment containing the lac promoter region was denatured in the absence and presence of CAP-cAMP. In 5 mM Na^+ a derivative melting curve of the 301 bp DNA alone shows two cooperatively melting peaks at 61.9°C and 66.4°C. The presence of CAP-cAMP resulted in the appearance of an additional small melting peak with a T_m of 64.4°C. This result suggested that site specific binding of CAP-cAMP stabilizes rather than destabilizes its binding region. The interpretation of the results with the 301 bp DNA are complicated however by the uncertain temperature stability of CAP, and the presence of many non-specific DNA sites and a secondary CAP binding site. Jensen and Von Hippel (14) have documented an example of a protein, ribonuclease A, which acts as a DNA melting protein at low temperatures, and becomes a DNA stabilizing protein after being heated to 55-60°C.

In order to help elucidate the thermodynamic binding properties of native CAP, we examined non-specific and site specific CAP binding to short DNA fragments. A 62 bp DNA containing the primary lactose promoter CAP site was employed. Figure 1 illustrates this fragment and the location of the CAP site. A 234 bp Hae III restriction fragment from pBR322 and a 144 bp DNA from the lactose promoter region were also examined. The influence of solvent, temperature and the presence of DNA on the binding activity of CAP was evaluated using a gel electrophoresis assay. Melting curve experiments were carried out in a solvent in which the transition midpoint temperature, T_m, of the 62 bp DNA was below the denaturation temperature of free CAP-cAMP. The melting curve studies indicated that site specific CAP-cAMP binding increases the stability of the 62 bp DNA by 13.3°C under conditions where non-specific binding stabilizes this DNA by 2-3°C. These results are only consistent with CAP-cAMP being a DNA stabilizing protein in its site specific binding mode.

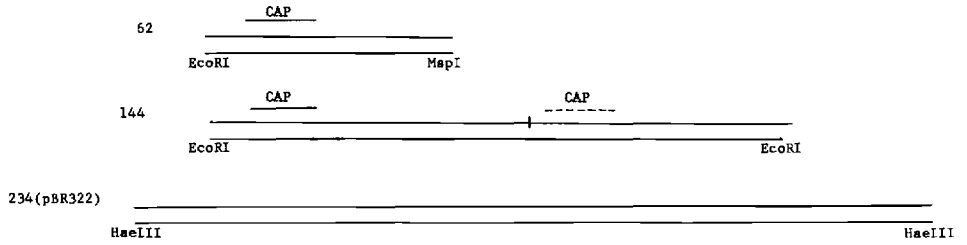


Figure 1. DNA restriction fragments used in this study. CAP sites are indicated for the 62 bp and 144 bp DNAs. Solid line underlining 'CAP' designates the primary functional site. The dashed line corresponds to a secondary CAP site located to the right of the major transcriptional startpoint (vertical slash).

MATERIALS AND METHODS

CAP isolation

CAP was isolated from an E. coli strain containing the plasmid pHA7 (16). This plasmid was constructed from an in vitro recombination of the plasmid pBR322 and the crp gene of E. coli. It was graciously provided by Hiroji Aiba. CAP was purified using a modification of the method of Boone and Wilcox (17). This method involved four steps. In step 1 a crude cell-free extract was prepared from 70 gm of frozen cells containing pHA7. A 10 ℓ. growth of pHA7 cells was disrupted in a French pressure cell at 10,000 psi followed immediately by centrifugation of the extract at 30,000 rpm for 90 min in a Beckman Ti-42.1 rotor. In step 2 the supernatant from step 1 was loaded onto a phosphocellulose column. We modified this step by employing a batch loading procedure.

Preequilibrated phosphocellulose was mixed with supernatant from step 1 by gentle stirring for 20 min. This mixture was allowed to settle and the top most layer of supernatant was removed. The phosphocellulose slurry was then loaded into the column. The remainder of this step was carried out as described by Boone and Wilcox (17). Active fractions from step 2, as measured by cAMP binding, were pooled and loaded onto a hydroxyappatite column (step 3). Active fractions from this column were collected, dialysed twice with 2 ℓ. of buffer C (20mM KPO_4, pH = 6.5, 1mM EDTA, 0.1 M NaCl) and loaded onto the final column. In this fourth step we substituted a denatured DNA-cellulose column for a DNA-DEAE-cellulose column. CAP was eluted using a pH jump from 6.5 to 8.0. Two 5 ml fractions were found to contain CAP at about 2 mg/ml each.

CAP concentrations were based on absorbance measurements using an

extinction coefficient of $\varepsilon_{280} = 3.5 \times 10^4\ M^{-1}\ cm^{-1}$, and confirmed using the cAMP binding assay (18). CAP was judged to be pure based on results from a 10% polyacrylamide gel containing SDS. The position of the isolated CAP in the gel was identical to a sample generously provided by A. Revzin. A single band was seen with 200 ug of protein loaded on the gel. CAP's ability to specifically bind to its DNA site was measured by the polyacrylamide gel electrophoresis assay of Garner and Revzin (19). About 25% of the CAP dimers was active for site specific binding in our preparation. This was similar to the value commonly found by others (19,20). CAP was stored in buffer C at 4°C, and at -20°C in 0.2M NaCl, 0.02M Tris (pH8 at 22°C), 0.1mM EDTA with 50% glycerol. Under both conditions there was no loss of activity over several months.

DNA Fragments

The 234 bp fragment was purified from Hae III digested pBR322 and was a generous gift of A. S. Benight. The 144 bp DNA fragment from the lactose promoter and the 62 bp restriction fragment obtained from it, were isolated from the plasmid pRMW27 (21) which has two copies of the 144 bp lac DNA cloned into the Eco R1 site of pVH51 DNA. The plasmid was maintained and grown in the E. coli strain M0. Purified plasmid was obtained from a modification of published procedures of Hardies and Wells (22). Following chloramphenicol amplified cell growth, lysis was carried out with lysozyme, Brij-58 and sodium deoxycholate. This procedure and the ensuing treatment with ribonuclease and phenol were similar to the published procedures. After phenol was removed by ether extraction and dialysis, the plasmid DNA was centrifuged to equilibrium in an ethidium bromide-CsCl gradient. The supercoiled DNA band was removed and the ethidium bromide extracted with isoamyl alcohol. The DNA was then dissolved and centrifuged in plain CsCl (ρ = 1.710 gm/ml). This second equilibrium centrifugation in CsCl alone significantly optimized the Eco RI digestion of the plasmid DNA. Ethidium bromide-CsCl runs did not remove some uncharacterized Eco RI inhibitor. The yield of pure plasmid varied from 2-3 mg/liter of cell growth.

The purified plasmid was cleaved with Eco RI. The 144 bp DNA was separated from the vector DNA by selective polyethylene glycol (PEG) precipitation in the presence of gelatin. The PEG step resulted in a considerable reduction of contaminating vector DNA. The lac fragment and the remaining vector were phenol extracted and separated by RPC-5 column chromatography. 7% polyacrylamide gel electrophoresis of 3 ug of the 144 bp DNA verified that it was purified to homogeniety.

The 62 bp DNA was prepared by cleaving the 144 bp DNA with Msp I and separating it from the other fragment by RPC-5 chromatography. The cleaved 144 bp DNA was phenol extracted prior to loading it on the RPC-5 column. Polyacrylamide gel electrophoresis was employed to verify the clean separation of the two DNAs indicated by the uv-absorbance elution profile. From 2 mgs of the 144 bp DNA, we recovered 0.74 mgs of the 62 bp DNA and 0.9 mgs of the 80 bp DNA.

Sample Preparation

Most of the melting curves were obtained using a solvent containing 1mM NaCl, 2mM Tris (pH = 8.0), 0.06mM EDTA and 40% dimethyl sulfoxide, DMSO. The DMSO was obtained from Aldrich Chem. Co. as 'ultrapure' spectrophotometric grade. When present, cAMP (Sigma Chem. Co.) was at a concentration of 10uM. Samples were prepared by first dialysing the DNA into 1.7mM NaCl, 3.3mM Tris, and 0.1mM EDTA. This solution was placed in ice, and DMSO was added in 20 ul amounts. Serial additions were continued until a final concentration of 40% DMSO was reached. If DNA solutions were not cooled before DMSO was added irreproducible results were obtained. This was due to heat generated from the exothermic reaction of DMSO and water, which apparently melted portions of the DNA. For samples with CAP, the CAP was added to the DNA in a solvent containing 0.1M NaCl and 3mM Tris. The mixture was then dialysed to the low salt solution and DMSO added as before. When cAMP was present it was added prior to the DMSO additions. A few experiments were carried out using 5mM sodium cacodylate (pH = 7.0) or 0.1M KCl + 0.04M Tris (pH = 8.0). CAP was dialysed directly into these solvents with or without DNA.

Thermal Denaturation

Absorbance-temperature profiles were measured at 268 nm using a Beckman Acta MVI spectrophotometer. Sample volumes were 0.7 ml. They were heated at a rate of about 7^{o}C/hour using a Lauda circulating bath. Further details of the experimental system have been previously given (23). The digital absorbance vs. temperature data was smoothed using the Savitsky and Golay method (24). Derivative melting curves were obtained by normalizing the absorbance change from 0 to 1.0, and taking the derivative of the smoothed melting curve. $d\theta/dT$, the change in the fraction of broken base pairs was plotted versus temperature. The T_m of a derivative melting curve was defined as the temperature at which $d\theta/dT$ was a maximum. Two or three transitions were taken to verify reproducibility.

Binding Assay

An assay involving polyacrylamide gel electrophoresis was employed to

determine the effect of solvent, temperature, and the presence of DNA on site specific CAP binding. Previous studies have demonstrated that CAP-cAMP lowers the mobility of short DNA fragments which contain specific CAP sites (19,20). 20 ul samples were taken from a solution of 0.5-5 uM CAP in a given solvent and at a specified temperature. In some experiments DNA was also present in the solution. After cooling the sample to 25°C, 2-3 ul of the 62 bp. DNA fragment in a similar solvent was added to give a final concentration ratio of one 62 bp. DNA molecule per four CAP dimers. This mixture, at a total volume of about 25 ul., was incubated for 15 minutes at room temperature, and then loaded onto a 7.5% polyacrylamide gel (46:1, acrylamide: bisacrylamide). Generally 3 ul of 0.3% bromphenol blue solution with 30% glycerol was added as a visual marker just prior to loading the samples. Omitting this dye solution gave identical results. Ethidium bromide staining was employed to visualize the DNA bands in the gel. Some experiments involved ^{32}P labeled 144 bp DNA as the probe (generously supplied by D. Dripps). For these experiments, autoradiography was employed to visualize DNA bands. The electrophoresis buffer contained 0.089M Tris, (pH8.5), 0.089M sodium borate, 0.1mM EDTA, and 10uM cAMP. Determinations of CAP thermal stability were made employing heating conditions identical to those used for the DNA melting curves. In some instances CAP stability measurements were made from the same CAP-DNA solutions being monitored for DNA melting. The results were the same for both cases.

RESULTS

CAP Binding to DNA in DMSO

An assessment of the effect of CAP-cAMP on DNA melting requires that CAP maintain its native binding activity through the melting region. We therefore sought a solvent in which CAP remains active for site specific binding at a temperature well above the T_m of the DNA alone. After investigating several solvents, it was determined that the low ionic strength solvent with 40% DMSO met these criteria. In this solvent, the 62 bp DNA alone has a melting temperature of 31.7°C (see below). Figure 2 shows the temperature stability of free CAP-cAMP for site specific binding. CAP-cAMP was removed from the 40% DMSO solvent at various temperatures, allowed to cool to 25°C, and challenged with the 62 bp DNA to determine whether it was irreversibly denatured. Below 41°C, CAP-cAMP retained its ability to alter the mobility of the probe DNA. The change in mobility was the same as that observed for CAP in the commonly used 100mM KCl solvent (not shown). Above

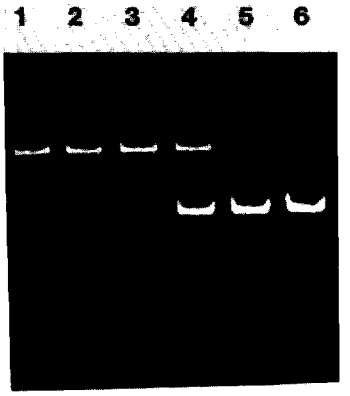


Figure 2. Gel assay of DNA specific binding after heating CAP-cAMP to various temperatures in 40% DMSO solvent. 10uM cAMP was also present. CAP concentration was 4.3 x 10^{-7}M. 20 ul of CAP was removed and 62 bp DNA added to generate 4/1 ratio of CAP/DNA (See Binding Assay). Lanes from left to right in ^{o}C; 25,30, 35,41,43,45. Lower band is free 62 bp DNA,upper band corresponds to CAP-DNA complex.

41^{o}C, the site specific binding of CAP-cAMP for the 62 bp DNA immediately decreases. It was also demonstrated that cAMP is required for site specific binding of CAP in 40% DMSO. CAP at a concentration of 0.6uM had no effect on the mobility of 0.15uM of the 62 bp DNA. Adding 10uM cAMP to an equivalent sample resulted in the DNA moving with the mobility characteristic of the site specific complex. Titration experiments showed that the same ratio of CAP to DNA was required to eliminate the free DNA band in 40% DMSO as in the 0.1M KCl solvent. Thus CAP shows site specific binding in the 40% DMSO solvent at a temperature 10^{o}C higher than the T_m of the 62 bp DNA.

The irreversible denaturation of CAP above 41^{o}C was correlated with a marked increase in light scattering from CAP-cAMP solutions. The apparent absorbance increased by 3-4 fold at 280 nm. This correlation did not occur when DNA was present either in the DMSO solvent or purely aqueous solvents. This is illustrated for a 5mM Na^{+} solvent in which CAP alone denatured around 48^{o}C in agreement with Takahashi et al. (25). When DNA was added CAP stability increased. Figure 3 illustrates the temperature stability of CAP-cAMP in a 5mM Na^{+} solution containing the 144 bp DNA. Below the midpoint of melting of the CAP-DNA complex (58^{o}C), CAP-cAMP retains its ability to bind in a site specific manner (Fig. 3a). Above 58^{o}C the 144 bp DNA is partially melted (Fig. 3b) and CAP-cAMP binding decreases. However no increase in light

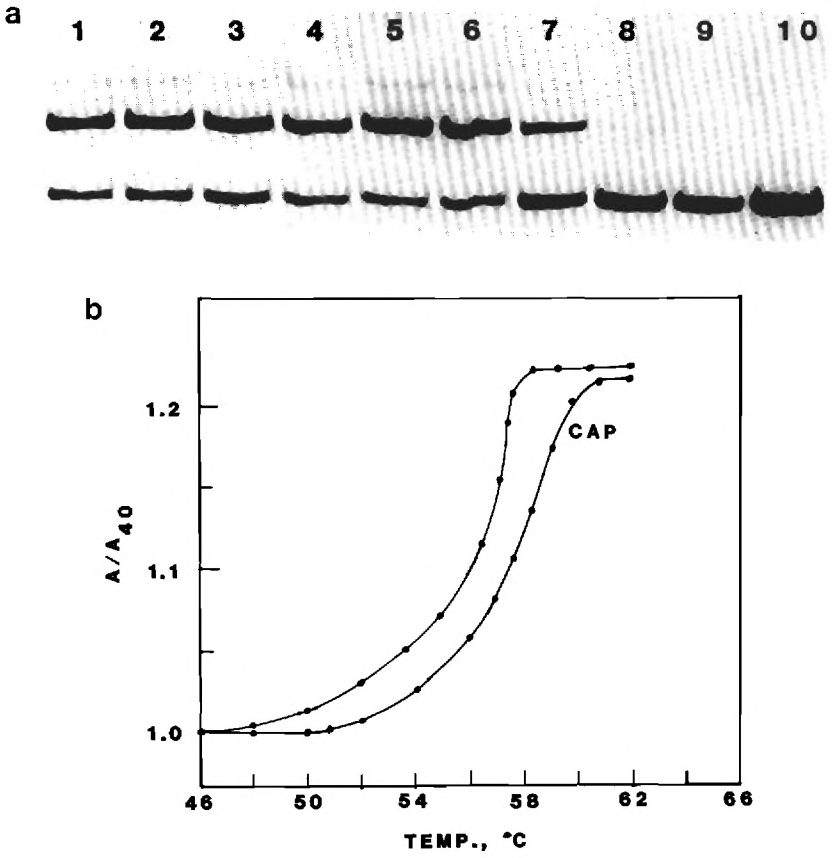


Figure 3. (a) Assay monitoring thermal stability of CAP-cAMP in the presence of 144 bp DNA in 5mM Na^+ . CAP/DNA ratio was 4/1 before adding the probe DNA. Aliquots were removed at specified temperatures and mixed with ^{32}P labelled unheated 144 bp DNA. This DNA was of low specific activity,and changed the CAP/DNA ratio to 3.2/1. This explains the presence of the free DNA band at low temperatures. Lanes 1,2,... 10 correspond to temperatures (^oC); 33.,35.5,41.,45.,51., 55.7,58.4,60.1,66.,70. Bottom band is free DNA. Faint top band is DNA with two bound CAP. (b) Melting curves of 0.3uM 144 bp DNA alone (—•—•—) and with 4/1 ratio of CAP/DNA (labeled CAP). Solvent was 5mM Na^+ and contained 10uM cAMP. The absorbance relative to the 40^{o}C value is plotted vs temperature.

scattering was observed even after heating the CAP-cAMP-DNA solution to 90^{o}C. Thus the absence of light scattering does not necessarily imply that CAP retains its native binding affinity.

The Influence of Non-Specific CAP Binding on DNA Stability

The effect of non-specific CAP binding on DNA stability was examined in 40% DMSO with the 62 bp lac DNA and the 234 bp DNA from pBR322. The 234 bp

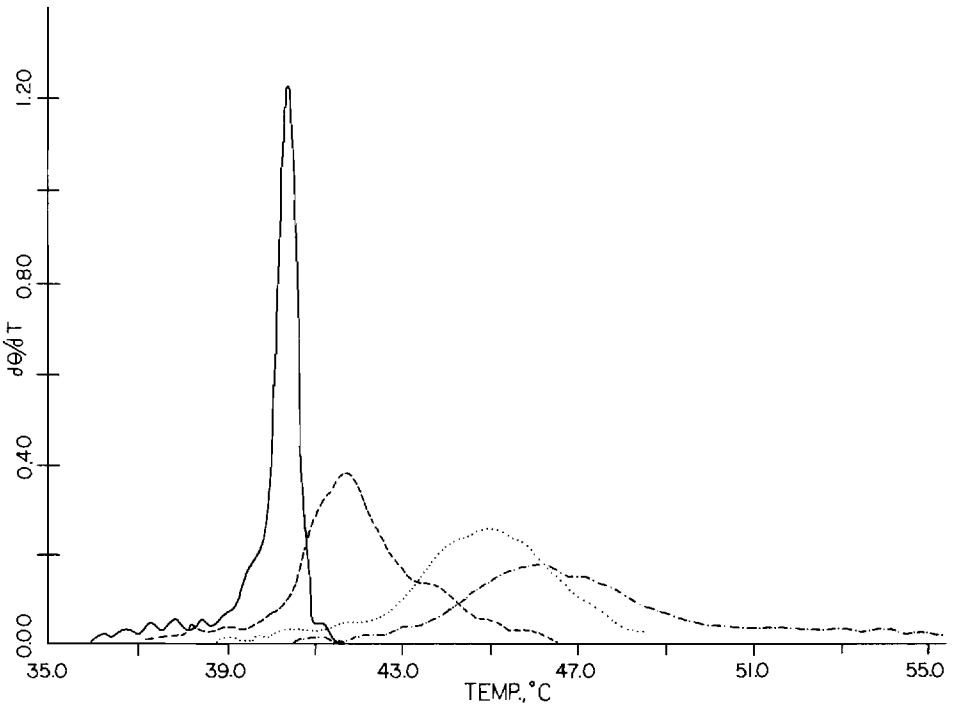


Figure 4. Derivative melting curves of 234 bp DNA in 40% DMSO solvent. Solid line has no CAP. Dashed line has a 7/1 ratio of CAP/DNA, dotted line has a 14/1 ratio of CAP/DNA, and dot-dash line has a 21/1 ratio of CAP/DNA. 10uM cAMP was in the solvent.

DNA does not have strong CAP-cAMP binding sites as indicated by the electrophoresis assay. Figure 4 shows derivative melting curves of the 234 bp DNA at several ratios of CAP-cAMP to DNA. Without CAP, the 234 bp DNA has a T_m of 40.4°C (± cAMP). Increasing amounts of CAP lead to a progressive increase in the transition temperature, a decrease in peak height and an increase in width of the transition. At 21 CAP-cAMP added per DNA the transition midpoint is increased by 5.4°C to 45.8°C and the width is increased by 5.6°C (full width at 0.5 maximum height).

Since the melting temperature of the 234 bp DNA saturated with CAP-cAMP is above the CAP denaturation temperature (41°C), some inactivation of CAP may occur prior to DNA melting. This was examined using the electrophoresis assay. Aliquots of the CAP-234 bp DNA solution were heated to various temperatures, cooled to 25°C, and then added to the 62 bp DNA. Site specific binding was observed until 45-46°C. Thus CAP bound to the 234 bp DNA was not

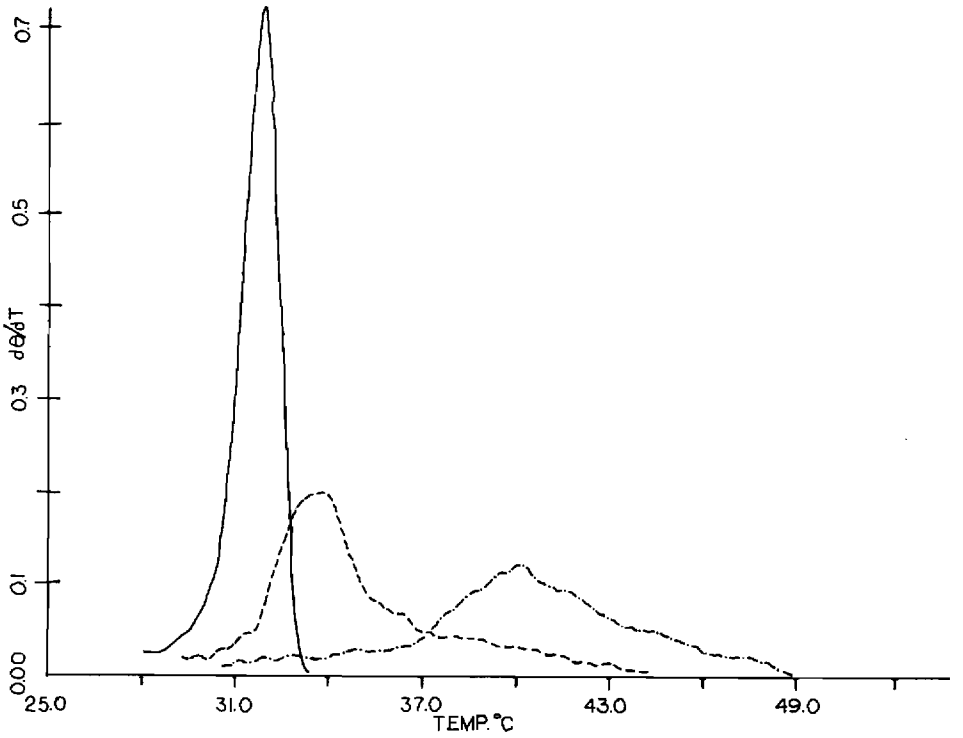


Figure 5. Derivative melting curves of 62 bp DNA in 40% DMSO solvent. Solid line has no CAP, dashed line has a 2.6/1 ratio of CAP/DNA, and dash-dot line has a 8/1 ratio of CAP/DNA.

irreversibly denatured up to the melting region.

The effect of non-specific CAP binding on the stability of the 62 bp lac DNA was also examined. In the absence of cAMP, CAP does not show site specific binding. Figure 5 shows derivative melting curves for the 62 bp DNA with increasing concentrations of CAP. Table I summarizes these results.

The effect of CAP alone on the 62 bp fragment melting curve was similar to that observed for the 234 bp DNA. The T_m of the transition increased with increasing CAP, and the transition cooperativity was reduced. Fewer CAP molecules were needed to saturate the shorter 62 bp DNA. The melting temperature of the 62 bp DNA fragment was increased to 40.1°C with eight CAP per DNA molecule. Since this DNA transition occurred below the temperature of CAP denaturation, the 8.4°C increment is not subject to uncertainties of CAP inactivation. Unlike poly(dA-dT), there is little or no indication of a biphasic transition at low CAP/DNA ratios (25). This can be explained by the

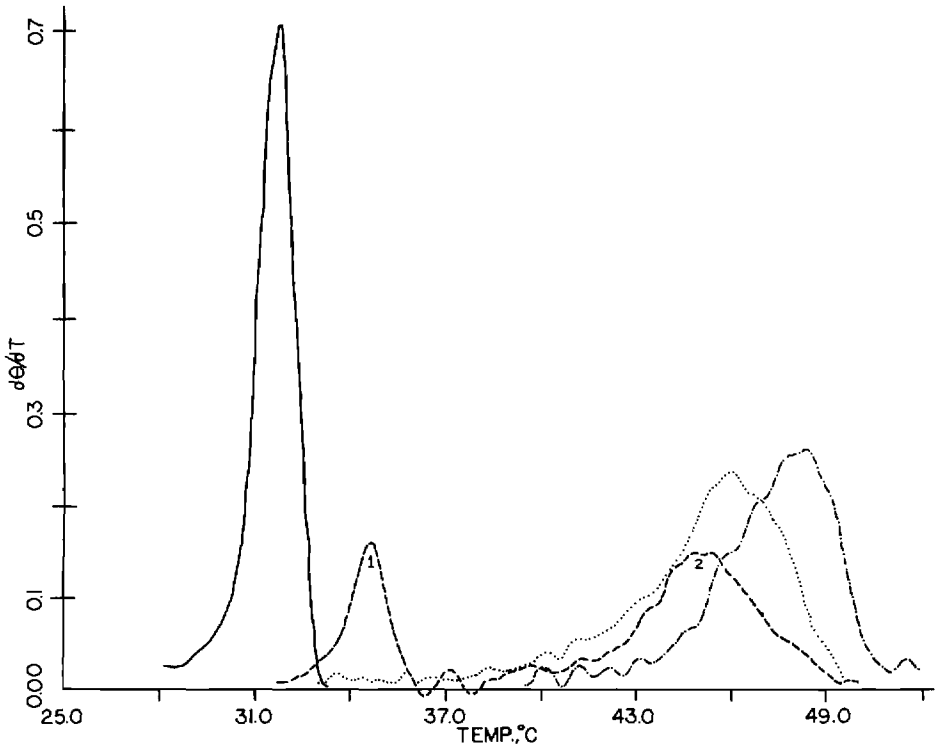


Figure 6. Derivative melting curve of 62 bp DNA with and without CAP-cAMP in 40% DMSO solvent. 62 bp DNA alone is solid line.Dashed line has 2.8/1 of CAP-cAMP/DNA,dotted line has 5.3/1 of CAP-cAMP/DNA,and dash-dot line has 8/1 of CAP-cAMP/DNA. 10uM cAMP was in the solvent.

short lengths of the DNAs, which are comparable to the cooperativity length of DNA melting. Since CAP binds strongly to DNA ends (27), it will initially bind non-cooperatively to all DNAs increasing their overall stabilities.

The effect of specific CAP binding on the 62 bp DNA and 144 bp DNA

The results described earlier demonstrated that CAP with cAMP exhibits site specific binding to the 62 bp DNA. Figure 6 shows the influence of CAP-cAMP on the melting transition of the 62 bp fragment in 40% DMSO. In the absence of CAP, the 62 bp DNA has a T_m of 31.7°C. The addition of 2.8 CAP dimers per DNA with cAMP results in a biphasic melting curve. The first subtransition has its peak at 34.7°C and a width of 1.6°C. The second subtransition has a peak at 45°C and a width of 4.1°C. The integrated areas of subtransition 1 vs. subtransition 2 are in the ratio of 0.3 to 0.7. The

TABLE I
Summary of results from derivative melting curves shown in figures 4, 5,and 6. W is the width of the transition or subtransition at half the maximum peak height.

DNA	cAMP	[CAP]/[DNA]	T , C	W , C
234 bp	+	0	40.4	.9
	+	7	41.8	3.4
	+	14	45.0	5.5
	+	21	45.8	6.5
62 bp	-	0	31.7	1.6
	-	2.6	33.8	3.0
	-	8.0	40.2	5.8
62 bp	+	0	31.7	1.6
	+	2.8	34.7,45.0	1.6,5.1
	+	5.3	46.0	5.3
	+	8.0	48.1	5.0

simplest interpretation of this ratio is that 70% of the 62 bp DNAs were stabilized 13.3°C, and 30% of the DNAs were stabilized 3°C. An analysis of the CAP-cAMP-DNA complexes formed at 2.8 CAP dimers per DNA indicates that subtransition 2 corresponds to site specific binding of CAP-cAMP and subtransition 1 corresponds to non-specific binding. Measurements discussed earlier showed that only 25% of CAP dimers are capable of site specific binding. Thus for 2.8 CAP dimers per DNA there are 0.7 CAP dimers per DNA capable of site specific binding. Assuming an equilibrium binding constant in 40% DMSO similar to that measured in other low ionic strength solvents (10^{11} M^{-1}), one can expect all of the CAP capable of site specific binding to be bound. Site specific binding can account for shifting 70% of the DNA to the higher melting subtransition. The remaining CAP can bind non-specifically to all DNA fragments. The 3°C increase observed for the first subtransition is similar to the 2.1°C increase observed for the non-specific binding of 2.6 CAP dimer per 62 bp DNA. (See Fig. 5 and Table I.)

Higher ratios of CAP-cAMP dimer to 62 bp DNA resulted in monophasic melting transitions (Fig. 6). The integrated areas of the transitions were the same as the combination of the two peaks observed at 2.8 CAP dimers/DNA. At 5.3 CAP dimers per DNA, the T_m is increased by 14.3°C. It is further increased to 48.1°C with 8 CAP dimers per DNA. The absence of a low temperature subtransition at CAP levels which saturate the specific DNA sites, and the equivalence in the transition areas are consistent with the earlier subtransition assignments. The electrophoresis assay verified that CAP-cAMP

retained its specific binding activity up to the melting region.

The influence of CAP-cAMP on the stability of the 144 bp DNA was also examined in 40% DMSO. In addition to the primary CAP site, the 144 bp DNA has a second site with a binding constant about 40 fold weaker than the primary site (26). The results of this study were similar to Fig. 6 (not shown). The 144 bp DNA alone melted with a T_m of 34.9°C. With one or two CAP dimers per DNA two melting peaks were observed at about 37-38°C and 42°C. The ratio of four CAP dimers per DNA gave one transition at 42.2°C.

DISCUSSION

The melting curve results show that CAP-cAMP stabilizes duplex DNA in both non-specific and site specific complexes. Saturation of the 234 bp DNA increased the melting temperature by 5.4°C. An 8.4°C increase was observed for the saturation of non-specific sites on the 62 bp DNA. The larger increase for the 62 bp DNA may be due to the more cooperative binding of CAP without cAMP (18), and/or the larger relative influence of the fragment ends. CAP has been observed to frequently bind to the ends of DNA fragments (27). The biphasic transitions observed at low ratios of CAP-cAMP to the 62 bp DNA (Fig. 6) and the 144 bp DNA can be explained by a combination of non-specific and site-specific binding. Only 25% of our CAP preparation was active for site-specific binding. When the number of active CAP molecules is less than the number of DNA molecules, a fraction of the DNA will interact with the inactive CAP. The low temperature subtransitions appear to be due to the inactive CAP molecules which are incapable of site-specific binding, but still capable of non-specific binding. The high temperature subtransition results from site-specific CAP-cAMP binding.

We theoretically asked the question, by how much would CAP-cAMP have to stabilize its 20 bp site in order to increase the 62 bp DNA melting curve by about 14°C? The helix-coil transition theory of DNA was employed (28,29). This calculation is not rigorously related to the experimental melting curves of the 62 bp DNA since a number of uncertain assumptions are necessary. DNA melting parameters are not known for the 40% DMSO solvent. Parameters determined in 0.1M NaCl were used. σ, the stacking cooperativity parameter was 4.5×10^{-5}, and $T_{GC} - T_{AT}$, the T_m difference between pure A.T and G.C DNAs was 43°C (29). The calculation also assumes that CAP-cAMP binding is approximately temperature independent. This latter assumption should be valid for temperatures below the free CAP denaturation temperature. Above the free CAP denaturation temperature, the situation is a complex one depending on the

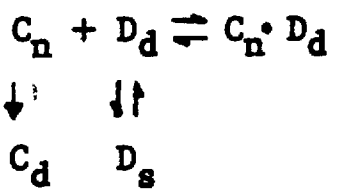


Figure 7. Major reactions governing the effect of CAP-cAMP on DNA melting. D_d and D_s correspond to duplex and single stranded DNA. C_n and C_d correspond to native (active) and denatured (inactive) CAP-cAMP. $C_n \cdot D_d$ correspond to the protein DNA complex.

rates of denaturation of free CAP, and the dissociation and reassociation rates of CAP to DNA (see below). Given these caveats it is of interest to note that increasing the melting temperatures of all base pairs in the region -50 to -69 by 35° C increased the T_m of the 62 bp DNA transition by 13.5°C. The predicted width of the transition broadened and the peak height decreased as was experimentally observed. Repeating this calculation for the 144 bp DNA resulted in an increase in T_m by about 7°C with similar changes in transition width and peak height. These theoretical results are in general accord with the melting curve data, and suggest that under the solvent conditions examined CAP-cAMP stabilizes its site by about 35°C.

Under physiological conditions it is likely that CAP-cAMP also acts as a DNA stabilizing protein. This is indicated by the results obtained by Kolb and Buc (30). CAP-cAMP when specifically bound to its lac site in supercoiled DNA in 100 mM KCl does not significantly unwind DNA. Although this does not establish CAP as a stabilizing protein, it undermines the notion that it is a melting protein. A direct assessment of CAP-cAMP's effect on DNA stability cannot unfortunately be made in 0.1-0.2M Na^+. If free CAP's denaturation temperature, T_D, is above the free DNA melting transition, T_m, one can determine if CAP is a stabilizing or melting protein. When T_D is less than the free DNA T_m (which occurs for the 62 bp DNA in 0.1M Na^+) the results are difficult to interpret. Figure 7 illustrates the major reactions governing how CAP will influence DNA stability. The rates of CAP denaturation, CAP-DNA on and off rates, and base pair opening and closing at the CAP site all influence the outcome. Once the protein DNA complex is heated past T_D, the potential exists for CAP to denature when it dynamically dissociates from DNA. As the temperature increases, one can expect the rate constant for CAP denaturation to increase. Base pair opening will also increase and will deplete duplex DNA sites. Since CAP binds weakly to single stranded DNAs

(18), the amount of free CAP, susceptable to rapid denaturation, will rise. For a situation where $T_m > T_D$, CAP may become inactive as soon as base pair opening becomes significant. The extent to which CAP stabilizes a DNA will be limited.

This appears to be the case in the 5mM Na^+ solvent. The T_m of the free 144 bp DNA (56.4°C) is greater than the free CAP denaturation temperature (48°C). CAP does not denature as long as the DNA remains a duplex. However, when the temperature exceeds the T_m, the large value of $T-T_D$ results in a high denaturation rate of CAP. A 1.6°C increase in the DNA's thermal stability is observed. This is less than the 7.3°C increase obtained for the same CAP-DNA ratio in the 40% DMSO solvent. In the latter solvent the free DNA T_m is less than T_D. Another result tends to corroborate the above conclusion. We examined the effect of CAP-cAMP on the stability of poly d(G-C) poly d(G-C) in the 40% DMSO solvent. This DNA polymer has a T_m of 56°C in this solvent. Saturating levels of CAP-cAMP (1 CAP dimer/12 base pairs) had essentially no effect on the transition T_m. The protein remained active for site specific binding (using the gel assay) until the DNA polymer reached the melting region. This behavior is qualitatively expected. Since T_m-T_D is large (15°C) CAP denatured as soon as base pair opening began to rise significantly. This did not occur for the non-specific binding of CAP to the 62 bp and 234 bp DNAs. Here $T_m<T_D$, and the stabilizing effect of CAP was observed.

The results and analysis indicate that site-specific CAP-cAMP binding stabilizes its promoter site. This physical behavior is consistent with CAP acting as a repressor by blocking RNA polymerase from strong initiation sites, and as an activator by displacing RNA polymerase from poor initiation regions to good initiation regions. The influence of CAP-cAMP on RNA polymerase may also involve other interactions as a result of DNA bending and protein-protein contacts. However the idea of a transmitted change in DNA stability has no experimental support.

ACKNOWLEDGEMENTS

We wish to thank the National Institutes of Health for support of this work through grant GM-33543, and Audrey Ralston and Ellen Cook for typing this manuscript. We also thank D. Brown, D. Dripps, and C. McCampbell for helpful comments on the manuscript.

REFERENCES

1. de Crombugghe, B.,Busby, S., Buc, H. (1984) Science 224, 831-838.
2. Taniguchi, T., O'Neill, M. and de Crombugghe, B.(1979) Proc. Natl. Acad. Sci. (USA) 76, 5090-5094.

3. Dickson, R. C., Abelson, J., Johnson, P., Reznikoff, W. S. and Barnes, W. H. (1977) J. Mol. Biol. 111, 65-75.
4. Majors, J., Ph.D. Thesis (1977) Harvard University.
5. Schmitz, A. (1981) Nucleic Acids Res. 9, 277-292.
6. Dunn, T. and Schleif, R. (1984) J. Mol. Biol. 180, 201-204.
7. Aiba, H. (1983) Cell 32, 141-147.
8. Gilbert, W. (1976) in RNA Polymerase (Losick, R. and Chamberlin, M. Eds) pp. 193-205. Cold Spring Harbor Lab., Cold Spring Harbor, N.Y.
9. Dickson, R. C., Abelson, J., Barnes, W. M. and Reznikoff, W. S. (1975) Science 187, 27-35.
10. Burd, J. F., Wartell, R. M., Dodgson, J. B., and Wells, R. D. (1975) J. Biol. Chem. 250, 5109-5113.
11. Musso, R., DiLauro, R., Adhya, S., and deCrombugghe, B. (1977) Cell 12, 847-856.
12. Malan, T. P. and McClure, W. R. (1984) Cell 39, 173-180.
13. Yu, X. M. and Reznikoff, W. S. (1984) Nucleic Acids Res. 12, 5449-5463.
14. Unger, B., Clore, G. M., Gronenborn, A. M. and Hillen, W. (1983) EMBO J. 2, 289-294.
15. Jensen, D. E. and von Hippel, P. H. (1976) J. Biol. Chem. 251, 7198-7214.
16. Aiba, H., Fujimoto, S. and Ozaki, N. (1982) Nucleic Acids Res. 10, 1345-1361.
17. Boone, T. and Wilcox, G. (1978) Biochimica et Biophysica Acta 541, 528-534.
18. Saxe, S. A. and Revzin, A. (1979) Biochemistry 18, 255-263.
19. Garner, M. M. and Revzin A. (1981) Nucleic Acids Res. 9, 3047-3060.
20. Fried, M. G. and Crothers, D. M. (1983) Nucleic Acids Res. 11,141-158.
21. Wartell, R. M. and Reznikoff, W. S. (1980) Gene 9, 307-319.
22. Hardies, S. C. and Wells, R. D. (1979) Gene 7, 1-14.
23. Benight, A.S., Ph.D. Thesis (1983) Georgia Institute of Technology.
24. Savitsky, A. and Golay, M. J. (1964) Anal.Chem. 36, 1627-1639.
25. Takahashi, W., Gronenborn, A. M., Clore, G. M., Blazy, B. and Bandres, A. (1982) FEBS Letters 139, 37-40.
26. Kolb, A., Spassky, A., Chapon, C., Blazy, B. and Buc H. (1983) Nucleic Acids Res. 11, 7883-7852.
27. Gronenborn, A., Nermut, M., Eason, P. and Clore, G. M. (1984) J. Mol. Biol. 179, 751-757.
28. Wartell, R. M. and Benight, A. S. (1982) Biopolymers 21, 2069-2081.
29. Wartell, R. M. and Benight, A. S. (1985) Physics Reports, in press.

Reprinted from Biochemistry, 1986, *25*, 2664.

Characteristics and Variations of B-Type DNA Conformations in Solution: A Quantitative Analysis of Raman Band Intensities of Eight DNAs†

Roger M. Wartell* and Juan T. Harrell
Schools of Physics and Biology, Georgia Institute of Technology, Atlanta, Georgia 30332
Received September 16, 1985; Revised Manuscript Received December 16, 1985

ABSTRACT: Raman spectra were obtained from four bacterial DNAs varying in GC content and four periodic DNA polymers in 0.1 M NaCl at 25 °C. A curve fitting procedure was employed to quantify and compare Raman band characteristics (peak location, height, and width) from 400 to 1600 cm^{-1}. This procedure enabled us to determine the minimum number of Raman bands in regions with overlapping peaks. Quantitative comparison of the Raman bands of the eight DNAs provided several new results. All of the DNAs examined required bands near 809 (±7) and 835 (±5) cm^{-1} to accurately reproduce the experimental spectra. Since bands at these frequencies are associated with A-family and B-family conformations, respectively, this result indicates that all DNAs in solution have a mixture of conformations on the time scale of the Raman scattering process. Band characteristics in the 800–850-cm^{-1} region exhibited some dependence on CG content and base pair sequence. As previously noted by Thomas and Peticolas [Thomas, G. A., & Peticolas, W. L. (1983) *J. Am. Chem. Soc. 105*, 993], the poly[d(A)]·poly[d(T)] spectra were qualitatively distinct in this region. The A-family band is clearly observed at 816 cm^{-1}. The intensity of this band and that of the B-family band at 841 cm^{-1} were similar, however, to intensities in the natural DNA spectra. Three bands at 811, 823, and 841 cm^{-1} were required to reproduce the 800–850-cm^{-1} region of the poly[d(A-T)]·poly[d(A-T)] spectra. This may indicate the presence of three backbone conformations in this DNA polymer. Analysis of intensity vs. GC content for 42 Raman bands confirmed previous assignments of base and backbone vibrations and provided additional information on a number of bands.

Several investigations have shown that the structure of duplex DNA is influenced by its base pair sequence and environment. X-ray diffraction studies on DNA crystals have demonstrated three classes of DNA structures, A, B, and Z (Rich et al., 1981; Dickerson et al., 1982; Shakked et al., 1983). Small localized alterations were observed within each type or family of DNA structure (Wang et al., 1980; Drew

† This work was supported by the National Institutes of Health (GM33543).

0006-2960/86/0425-2664$01.50/0

et al., 1981; Milane et al., 1984). In solution, DNA is subject to a variety of internal motions and environmental influences that, in general, can not be recreated in the solid state. These factors can be expected to produce conformations that are more varied and potentially different from structures observed in the solid state. The dynamic character of DNA is indicated by fluctuational base pair opening (Englander & Kallenbach, 1983; Mirau & Kearns, 1985; Wartell & Benight, 1982) and topological, bending, and torsional considerations (Klysik et al., 1981; Hagerman, 1981; Shore & Baldwin, 1983; Horowitz & Wang, 1984). NMR studies showed that the same DNA molecule does not necessarily have the same average solution and solid-state structures (Reid et al., 1983; Sarma et al., 1985).

Raman spectroscopy is one of several techniques that can be employed to investigate DNA conformation in solution. It has the advantage of being applicable to both large and small nucleic acids. The complexity of DNA Raman spectra has necessitated that an empirical approach be employed in assigning Raman bands to specific molecular groups and conformations. Studies on model compounds (Lord & Thomas, 1967; Thomas & Kyogoku, 1977; Nishimura et al., 1978, 1984) and DNA fibers and crystals (Erfurth et al., 1975; Thomas & Peticolas, 1983a; Martin & Wartell, 1982; Thomas et al., 1984) have been of great value in establishing assignments of DNA Raman bands. The three structural classes of DNA, A, B, and Z, can be distinguished by their Raman vibrational spectra.

In order to establish a basis for examining possible sequence-dependent variations in DNA Raman bands, we examined the Raman spectra of eight DNAs in 0.1 M NaCl solution at 25 °C. Four of the DNAs were heterogeneous-sequence naturally occurring DNAs varying in percent GC from 26.5 to 72%. The other four DNAs were synthetic DNA polymers. A curve fitting procedure was applied to quantify Raman band characteristics (peak location, height, and width) from 400 to 1600 cm^{-1}. This allowed us to examine the effect of percent GC content and base pair sequence on the intensities of previously assigned base and backbone vibrations. We have confirmed many assignments and obtained additional information on several Raman bands.

This study also provided new results on several aspects of DNA conformation in solution. Evidence is given for the presence of C3′-endo and C2′-endo family backbone bands in all DNA spectra. The quantitative analysis of the synthetic DNAs' spectra revealed features relating to their conformation. Studies on these polymers have been previously reported (Pohl et al., 1973; Thomas & Peticolas, 1983b; Fidor et al., 1984; Thomas & Benevidis, 1985; Wartell et al., 1983).

Material and Methods

Samples. The four naturally occurring DNAs used in the study were purchased from Sigma Chemical Co. They were from the bacteria *Clostridium perfringens* (26.5% GC), *Bacillus subtilis* (44% GC), *Escherichia coli* (50% GC), and *Micrococcus luteus* (72% GC). They were extracted with phenol–chloroform and dialyzed into 10 mM NaCl + 0.1 mM ethylenediaminetetraacetic acid (EDTA). The DNAs were then precipitated with 2–3 volumes of ethanol at −70 °C, backwashed with 75% ethanol twice to remove excess salts, and lyophilized to dryness. Gellike solutions were prepared for Raman experiments by adding 0.1 M NaCl to give concentrations at about 40–80 mg/mL. The synthetic DNAs poly[d(A-T)]·poly[d(A-T)], poly[d(A)]·poly[d(T)], and poly [d(G-C)]·poly[d(G-C)] were purchased from P-L Biochemicals. They were prepared for Raman spectroscopy in a manner similar to that described above. All DNAs were characterized by their UV spectra and DNA denaturation curve in 0.1 M NaCl + 1 mM sodium phosphate (pH 7.) + 0.1 mM EDTA except for poly[d(G-C)]·poly[d(G-C)]. This DNA was melted in 5 mM Na^+. Poly[d(T-G)]·poly[d(C-A)] was generously provided by Dr. J. Alderfer.

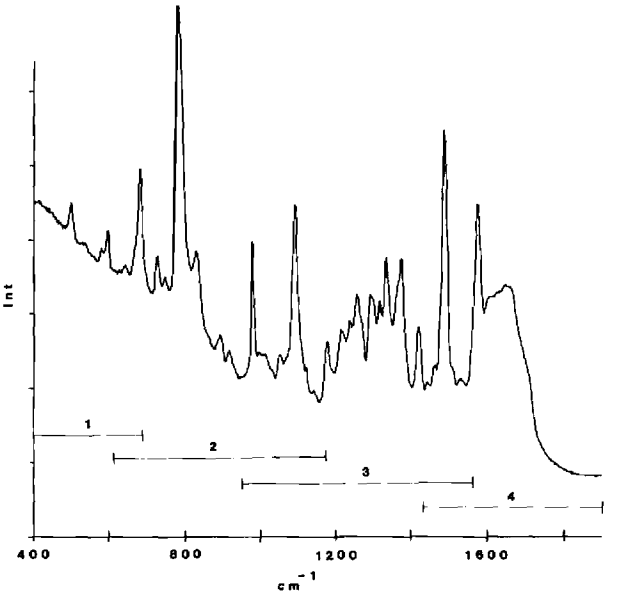


FIGURE 1: Raman spectra of *M. luteus* DNA is shown from 400 to 1850 cm^{-1}. The four regions designated below the spectra were the intervals employed in the quantitative analyses of the Raman band characteristics. The broad water band at 1645 cm^{-1} made the evaluation of nearby bands unreliable.

Raman Measurements. Spectra were obtained by using the 514.5-nm line of an argon ion laser. Laser power of 150–200 mW was employed to record spectra between 400 and 1900 cm^{-1}. The effective slit width of the spectrometer was 4.0 cm^{-1}. Scanning of the spectrometer (Spex 1401) and photon counting from the photomultiplier (RCA 31034) were coordinately controlled by a Tektronix 4051 computer. Sample volumes of 5–10 μL were placed in quartz capillary cells and held against an aluminum block whose temperature could be controlled by circulating water. All spectra were obtained at 25 °C.

Figure 1 shows a typical Raman spectrum from this study. It is the result of four consecutive scans with counts accumulated for 2 s/cm^{-1}. At least two separate spectra were taken of each sample. Band peak heights, widths, and locations were evaluated by a nonlinear least-squares curve fitting procedure. The procedure is based on the Marquardt (1964) algorithm and has been briefly described previously (Wartell et al., 1983). A polynomial up to order seven represented the background. Raman bands were reproduced by employing Lorentzian functions convoluted with the instrument function. For a given spectral region, the number of bands, their locations, and the order of the polynomial background curve were first estimated from the experimental spectra. Widths were initially estimated as 8 or 15 cm^{-1}. On the basis of this first guess, a linear least-squares fitting program was employed to provide a first-order estimate of peak heights and background parameters. The Marquardt algorithm then iteratively refined the fit by varying peak locations, widths, and heights as well as background parameters. The algorithm stopped when specified least-squares conditions were met or when 500 iterations were reached. A copy of the program in Fortran V appropriate for the Control Data Corp. Cyber 850 computer with IMSL programs is available upon request. The curve fitting proce-

dure was applied to the four regions designated in Figure 1. These regions were large enough to follow the background accurately without exceeding the memory required to fit the experimental spectra.

Shapes of the calculated background curves were essentially the same as those of the experimental solvent spectra. Not surprisingly, background curves of overlapping regions did not splice together perfectly. An estimate of the error generated by these end effects was made by comparing Raman band characteristics evaluated from Raman peaks shared by overlapping regions. Differences were generally no larger than repeated analyses of a single region. They were included in determining average Raman band characteristics. Each spectral region was analyzed 3 times, and the band characteristics were averaged. The second region for example required 18 bands from 640 to 1150 cm^{-1}. These bands were the minimum number required to reproduce this region in the spectra of the four natural DNAs. Band positions and widths were similar among the DNAs. In some cases, small GC-dependent frequency shifts were noted (see below). Similar results were obtained in the other spectral regions. Although some of the broader calculated bands may actually result from two vibrational bands, the analysis provides a minimum set of Raman bands, which enables an accurate intensity comparison to be made in regions of overlapping peaks.

Results

Raman Spectra of Natural DNAs. A comparison of Raman spectra of the natural DNAs was made in order to verify the percent GC dependence predicted by previous band assignments. Figure 2 shows spectra of *C. perfringens* DNA, *E. coli* DNA, and *M. luteus* DNA with backgrounds subtracted. The spectra were normalized to the 1093-cm^{-1} band associated with the PO_2^- symmetric stretching mode. The narrow Raman peak at 980 cm^{-1} was from 0.02 M sodium sulfate added to the solvent. This band or "line" provided an internal frequency standard. The effect of GC content on the intensities of a number of Raman peaks can be readily observed. Examples in the 400–1200-cm^{-1} region were the intensity decreases at 669 cm^{-1} and intensity increase at 682 cm^{-1} as the percent GC increases. These changes were expected since the 669-cm^{-1} band was assigned to thymine and the 682-cm^{-1} peak was assigned to a guanine–deoxyribose mode (Hartman et al., 1973). Similarly, the 730-cm^{-1} adenine vibration and 750-cm^{-1} thymine band decreased with increasing GC content, while the 782-cm^{-1} cytosine band increased. The peak around 833 cm^{-1} has been assigned to a phosphate–deoxyribose vibration characteristic of the B-DNA conformation (Erfurth et al., 1975). Its intensity did not significantly change with increasing percent GC.

In the 1000–1800-cm^{-1} region, the 1488-cm^{-1} band increased in intensity, the 1512 cm^{-1} band decreased, and the 1420-cm^{-1} band remained approximately constant. These observations correspond with previous assignments. The 1488-cm^{-1} mode is a purine ring vibration dominated by guanine, the 1512-cm^{-1} band is assigned to an adenine vibration, and the 1420-cm^{-1} band is assigned to deoxyribosyl CH_2 groups (Prescott et al., 1984). Changes in band intensities from 1150 to 1450 cm^{-1} were more difficult to assess. Observed peaks were the sum of contributions from overlapping Raman bands. This also occurred for the 750–840-cm^{-1} region and from 980 to 1150 cm^{-1}. In order to obtain a more quantitative comparison of Raman bands throughout the 400–1800-cm^{-1} region, we will turn to the results from the curve fitting analyses of both natural and synthetic DNAs.

Intensity vs. Percent GC of Raman Bands. Figure 3 plots

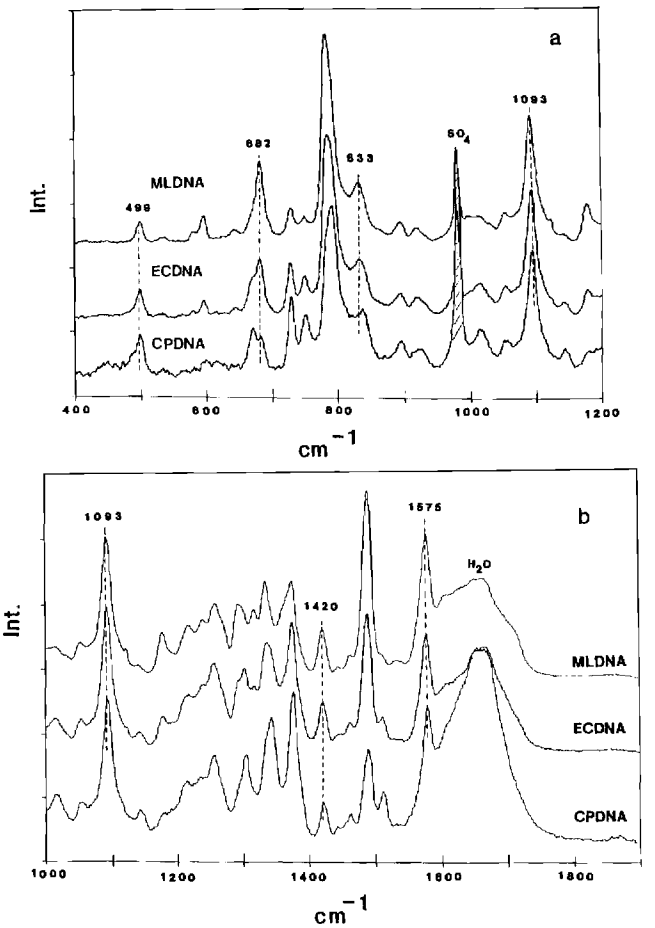


FIGURE 2: Comparison of Raman spectra of DNA from *C. perfringens* (CPDNA), *E. coli* (ECDNA), and *M. luteus* (MLDNA). Background curves were subtracted and the spectra normalized to the 1093-cm^{-1} band. (a) 400–1200-cm^{-1} region; (b) 1000–1900-cm^{-1} region.

the relative intensities of four bands from 400 to 700 cm^{-1} vs. percent GC. The filled and open symbols designate natural and synthetic DNAs, respectively. All four bands in Figure 3 have been associated with base ring vibrations. The 499-cm^{-1} band has been assigned to the sum of two vibrational modes due to guanine and thymine (Prescott et al., 1984). Our data are consistent with this assignment. Both AT DNA polymers and poly[d(G-C)]·poly[d(G-C)] have bands at 499 ± 1 cm^{-1}. In addition, the linear extrapolation of the natural DNA data indicate nonzero intensities at 0 and 100% GC. With the exception of poly[d(A-T)]·poly[d(A-T)], all the synthetic DNAs have intensity values at 499 cm^{-1} that fit the line drawn through the natural DNA data. The intensity vs. percent GC of the 596-cm^{-1} band in Figure 3 is consistent with this band being a guanine and/or cytosine vibration. Studies carried out with poly[d(I-C)]·poly[d(I-C)] and poly[d(G-meC)]·poly[d(G-meC)] indicate it is a cytosine vibration (D. Brown and R. M. Wartell, unpublished results). The GC dependence of the 669-cm^{-1} thymine band and 682-cm^{-1} guanine vibration is as anticipated. Except for slight deviations of the DNA polymers' intensities for some of the bands, the synthetic DNA data are in general accord with the linear extrapolations from the natural DNA data.

Figure 4 continues the intensity vs. percent GC plots for bands at 729, 750, 782, and 791 cm^{-1}. A small 729-cm^{-1} band for poly[d(G-C)]·poly[d(G-C)] indicates a small contribution at this frequency from a vibration not involving adenine. The 782- and 791-cm^{-1} bands form part of a convoluted set of bands from 760 to 840 cm^{-1}. This section has backbone bands

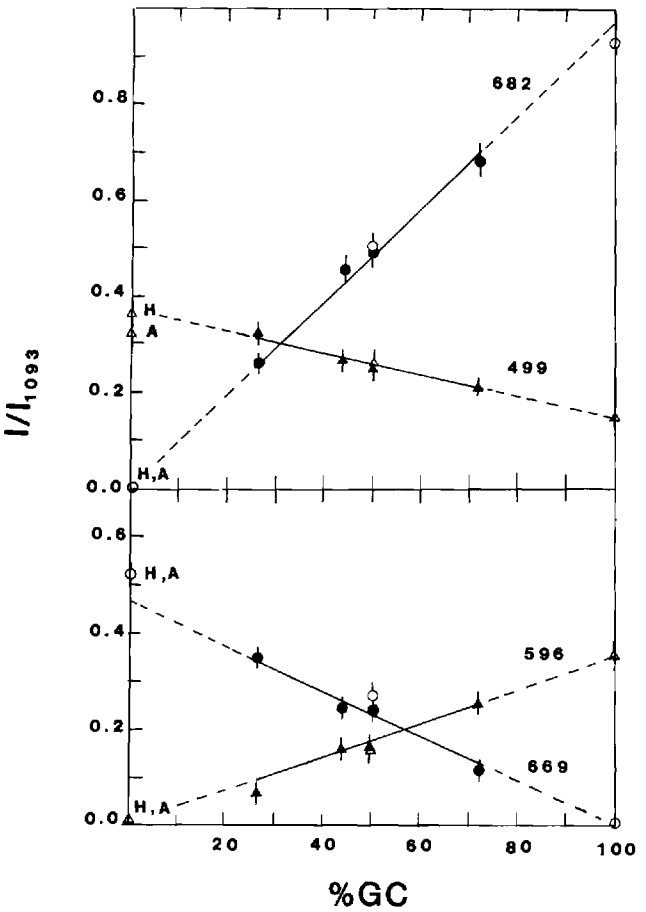


FIGURE 3: Raman intensities vs. percent GC are plotted for bands at 499, 596, 669, and 682 cm^{-1}. Filled symbols correspond to natural DNAs; open symbols are periodic DNAs. "H" refers to homopolymer AT DNA poly[d(A)]·poly[d(T)]; "A" refers to alternating AT DNA poly[d(A-T)]·poly[d(A-T)]. Intensities are relative to the 1093-cm^{-1} band.

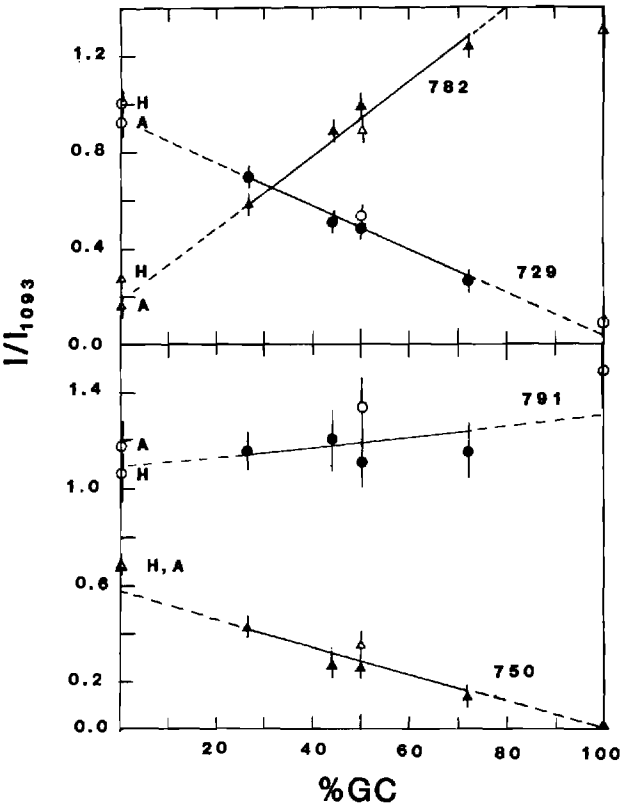


FIGURE 4: Raman intensities vs. percent GC for bands at 729, 750, 782, and 791 cm^{-1}. See Figure 3 legend for symbol description.

that show major intensity changes in the B ⇄ Z transitions. We have quantified these bands using both the 640–1150-cm^{-1} and 700–800-cm^{-1} regions. Results were similar in both cases. The 782-cm^{-1} band has been assigned to a cytosine vibration (Erfurth & Peticolas, 1975; Prescott et al., 1984). Figure 4 implies an additional contribution. Both AT DNAs require a small band around 782 cm^{-1} to fit experimental spectra (see Figures 6 and 7). Consistent with this finding is the extrapolated 0% GC end point deduced from the natural DNA data. The above results can be of significance in assigning calculated modes to observed Raman bands. Another somewhat surprising observation is the intensity for poly[d(G-C)]·poly[d(G-C)] at 782 cm^{-1}. It is lower than projected from the other DNA data. This result was observed with several independent spectra. It may reflect some unique conformational property for this DNA. The 791-cm^{-1} band was initially assigned to a backbone vibration by Peticolas and collaborators (Erfurth et al., 1975). Although the measurement error of this band's intensities is large, the data are consistent with this interpretation.

800-850-cm^{-1} Region. One of the unexpected observations of the study was the presence of a band around 809 cm^{-1} in the spectra of all the natural and synthetic DNAs. This band cannot be visually detected in spectra from the natural DNAs, poly[d(G-C)]·poly[d(G-C)], and poly[d(T-G)]·poly[d(C-A)]. However, the curve fitting analysis required it to accurately fit each spectra of these DNAs. Figure 5 demonstrates this for *M. luteus* DNA. Figure 5a shows the calculated fit when a band around 809 cm^{-1} was included. Figure 5b displays the result of the curve fitting analysis in the absence of this band. The fitting algorithm attempts to compensate for the absence of the 809-cm^{-1} band by increasing the widths of other bands to 40–50 cm^{-1}. Even with this compensating flexibility, the calculated spectra did not reproduce the experimental spectra. Similar results were obtained for the other DNAs (see paragraph at end of paper regarding supplementary material). We also note that six bands is the least number of bands that fit the data. The 809-cm^{-1} band of the natural DNAs was generally broader than that of the synthetic DNAs and could be fit with a larger number of modes. A Raman peak of high intensity near 809 cm^{-1} is characteristic of an A-type DNA backbone. It has been correlated by experiment (Erfurth et al., 1975) and theory (Lu et al., 1977) with a vibrational mode involving both the deoxyribose group in a C3′-endo ring pucker and the C5′–O–P(O_2^-) atoms.

Figures 6 and 7 display the 800-cm^{-1} region of the AT DNA polymers. The spectra of these DNAs differ from the other DNA spectra between 800 and 850 cm^{-1}. The "809"-cm^{-1} A-family band and "835"-cm^{-1} B-family band are located at 816 and 841 cm^{-1}, respectively, in poly[d(A)]·poly[d(T)]. The curve fitting analysis of poly[d(A-T)]·poly[d(A-T)] spectra shows a requirement for an additional Raman band at 823 cm^{-1} (Figure 7). This band has a width and peak height similar to its neighboring bands. It is not possible to accurately fit this DNA's Raman spectra with only two bands. This was the only DNA we examined that required three bands from 800 to 840 cm^{-1}. Interpretation of this band is postponed until later.

Figure 8 shows band intensity vs. percent GC for the 809-cm^{-1}, 835-cm^{-1} and two other backbone modes. The 816-cm^{-1} band of poly[d(A)]·poly[d(T)] and the 811-cm^{-1} band of poly[d(A-T)]·poly[d(A-T)] are considered as "809"-cm^{-1} bands. The change in intensity with percent GC is small for

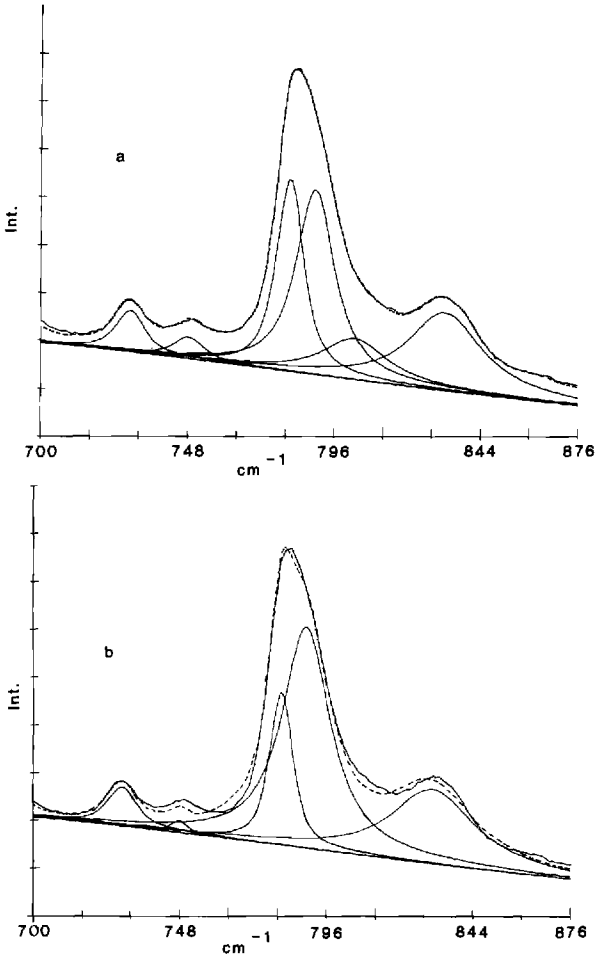


FIGURE 5: Experimental and fitted spectra of *M. luteus* DNA from 700 to 876 cm^{-1}. Solid lines denote experimental spectra, Raman bands, and background curve. Dashed line denotes the sum of background curve and Raman bands. (a) Six bands employed in analysis procedure. (b) Five bands used in analysis procedure. The same fitting criteria were used in both cases.

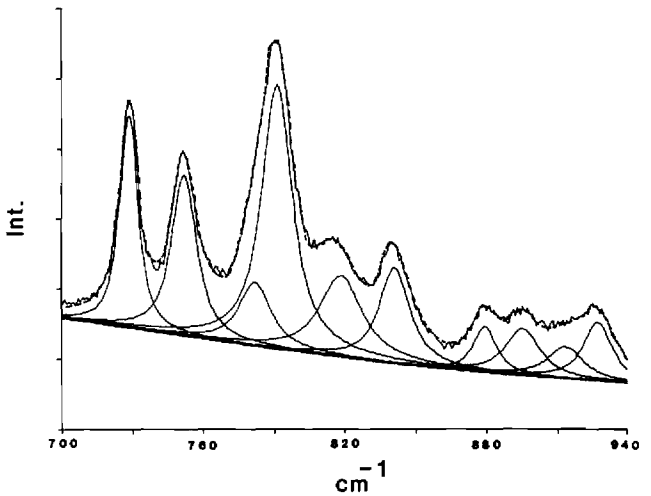


FIGURE 6: Raman spectra of poly[d(A)]·poly[d(T)] and Raman bands and background curve determined from the analysis procedure. Solid lines denote experimental spectra, background curve, and Raman bands. Dashed line denotes sum of background curve and Raman bands.

all of these bands. This is consistent with their being backbone modes. Deviations from expected intensities are observed for poly[d(G-C)]·poly[d(G-C)] at 835 cm^{-1} and for the AT and GC DNA polymers at 924 cm^{-1}.

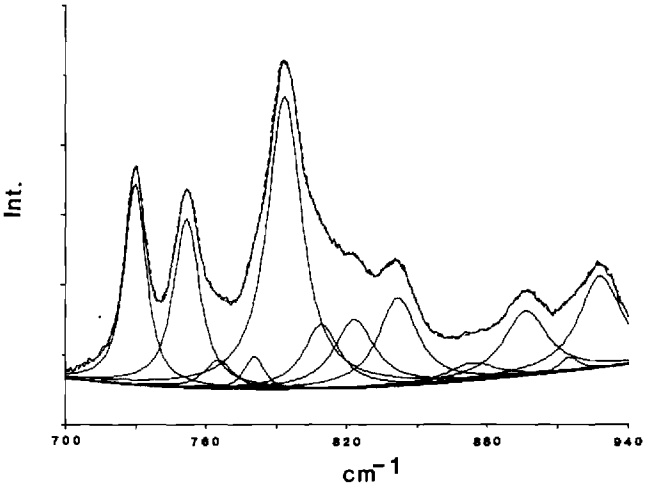


FIGURE 7: Raman spectra of poly[d(A-T)]·poly[d(A-T)] and Raman bands and background curve determined from the analysis procedures. Solid lines denote experimental spectra, background curve, and Raman bands. Dashed line denotes sum of background curve and Raman bands.

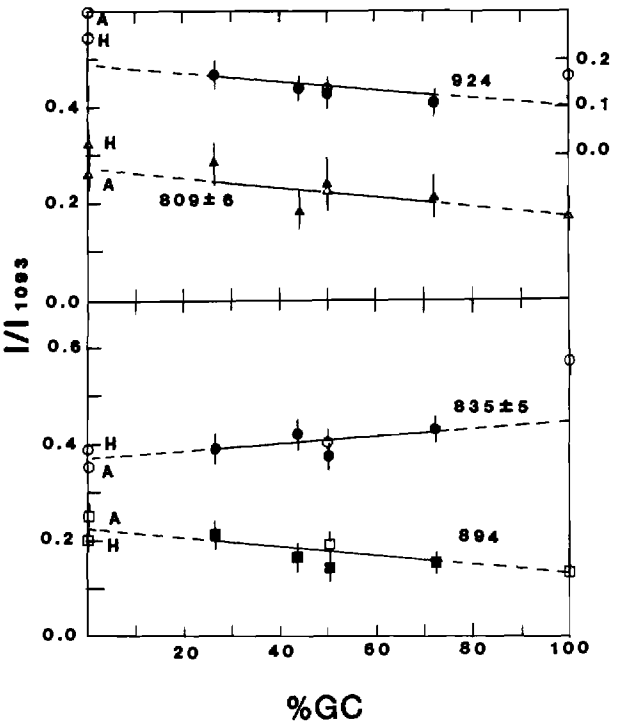


FIGURE 8: Raman intensities vs. percent GC for bands at 809, 835, 894, and 924 cm^{-1}. See Figure 3 legend for symbol description.

Intensity vs. percent GC plots were obtained for a number of bands in the 1000–1600-cm^{-1} region. This portion of the spectra has many overlapping bands. We reiterate that the bands required for this spectral region form a minimum set. Some are likely to correspond to overlapping modes. Although the percent GC dependence of these Raman bands are generally consistent with prior assignments, several exceptions are noted. Bands at or near 1052, 1213, 1237, 1256, 1333, and 1420 cm^{-1} show little change in intensity for all the DNAs. The 1052- and 1420-cm^{-1} bands were assigned to backbone modes (Prescott et al., 1984). The constant intensities of these bands with percent GC reinforce these assignments. The band found near 1213 cm^{-1} was broad and overlapped with the 1237-cm^{-1} band. Both bands were assigned to thymine ring modes by Prescott et al. (1984). A different assignment for the 1237-cm^{-1} peak (cytosine) was made by Chou and Thomas

Table I: Characteristics and Assignments of Raman Bands Observed in both Natural and Periodic DNAs[a]

nominal frequency (cm^{-1})[b]	width at half-height (cm^{-1})[c]	assignment	nominal frequency (cm^{-1})[b]	width at half-height (cm^{-1})[c]	assignment
490	14.0	Ade/Thy[d]	1115	25.0	
499	9.0	Gua, Thy	1143	10.0	Ade/Thy
534	8.0	Ade	1178	16.0	Gua/Cyt
579	7.0	Gua/Cyt	1192	7.0	
596	8.5	Cyt	1213*	28.0	bk/mm
641	11.0	Ade, Cyt	1237*	12–23	bk/mm
670	10.0	Thy	1256	19.0	Cyt, Ade
682.5	11.5	Gua-Rib	1268*	9–21	
729	8.0	Ade, u	1302	13.5	Ade
750	11.5	Thy	1317	12.0	Gua
782	8.0	Cyt, u	1333	14.0	bk/mm
791	14.5	bk	1342	16.0	Ade
809*	24.5	bk	1361	17.0	Gua/Cyt
835*	22.0	bk	1375	15.0	Ade, Gua, Thy
894	20.0	bk	1420	11.0	Rib
924*	19.0	bk	1444	6–20	
998	20.0		1460	10–30	
1013	27.0	Ade/Thy	1488	13.0	Gua, Ade
1052	12.0	bk	1512	7.0	Ade
1067*	26		1575	14	Gua, Ade
1093	16	bk (PO_2^-)			

[a] These bands were observed for most if not all DNAs. A number of bands specific for individual DNA polymers are not listed. [b] Bands that varied in frequency by more than ±2 cm^{-1} among the DNAs are starred marked with an asterisk (*). [c] If the SD of the band width was less than 25% of the average width, then the average value is given. Otherwise, the range is given. [d] A slant (/) means and/or. A comma implies and. Gua, guanine; Cyt, cytosine; Ade, adenine; Thy, thymine; u, unknown; bk, backbone; Rib, deoxyribose; mm, multiple base modes.

(1980) and Strommen and Peticolas (1982). Since the intensity of neither band changes with GC content, it is unlikely that they correspond to vibrations of a single base. A variation of ±6 cm^{-1} was observed for the peaks of these bands among the synthetic DNAs. These bands may result from a backbone mode and/or several base modes. Theory and experiment indicate an antisymmetric phosphate stretch in this region (Lu et al., 1977; Forrest & Lord, 1977).

Several bands assigned to adenine—at 1302, 1342, and 1512 cm^{-1}—show the expected decrease in intensity with increasing percent GC. Large differences occur in the intensities of these bands for poly[d(A)]·poly[d(T)] and poly[d(A-T)]·poly[d(A-T)]. This implies that these adenine modes are sensitive to conformational differences between these DNAs. Table I lists bands observed in the DNAs studied. Assignments are based on this and previous work.

Discussion

The requirement for a band around 809 cm^{-1} in all spectra indicates that a portion of the deoxyribose rings in "B-type DNA" assume a C3′-endo family conformation (throughout the Discussion, we will refer to this A-DNA-type deoxyribose–phosphate mode as a C3′-endo family mode). This conclusion is based on the assignment of the deconvoluted 809-cm^{-1} band to the same backbone mode that produces a large peak at this frequency for A-DNA fibers. The small change in band intensity vs. percent GC is consistent with this assignment. Also, bandwidths are similar to those found for the 835-cm^{-1} backbone band.

The coexistence of the 809- and 835-cm^{-1} bands can be due to the dynamic fluctuation of every deoxyribose ring and/or to two populations of furanose conformations, each one correlated with specific nucleotide sequences. Drew et al. (1981) observed a distribution of furanose conformations in B-DNA oligomer crystals, but with the exception of a guanine base at the 3′ end, only 1 out of 20 internal bases had a furanose pucker close to C3′-endo. NMR studies of various DNA sequences have not yet indicated a population of C3′-endo-type sugar rings (Reid et al., 1983; Assa-Munt & Kearns, 1984; Sarma et al., 1985). These results suggest that the Raman

Table II: Frequencies of Bands in the 800–850-cm^{-1} Region

DNA	C3′-endo family (cm^{-1})	C2′-endo family (cm^{-1})	other band (cm^{-1})
poly[d(A)]·poly[d(T)]	816	841	
poly[d(A-T)]·poly[d(A-T)]	810	842	823
C. perfringens	811	837	
B. subtilis	808	834	
E. coli	805	834	
poly[d(T-G)]·poly[d(C-A)]	807	837	
M. luteus	803	833	
poly[d(G-C)]·poly[d(G-C)]	803	831	

experiments are sampling the dynamic character of all rings, although a small equilibrium population of C3′-endo rings is likely. The intensity of the 809-cm^{-1} band determined from this study as well as from calf thymus DNA in solution is about 20% of the 809-cm^{-1} band in A-DNA calf thymus fibers (Martin & Wartell, 1982). This suggests that 20% of the deoxyribose rings of solution DNA are in the C3′-endo family conformation. The remainder are presumably in the C2′-endo family.

Another observation that is probably related to conformational variability is the influence of percent GC on the frequencies of the nominal 809- and 835-cm^{-1} bands. This has also been noted by others (Thomas & Peticolas, 1983b; Thomas & Benevidis, 1985). Table II shows an upward shift of about 10 cm^{-1} for both bands with increasing percent GC. One explanation for this behavior is that these "backbone" modes actually involve base atom motions (Lu et al., 1979) and are sensitive to base pair type. Alternatively, it may reflect small differences in backbone conformation with percent GC. Evidence for a relationship between the frequency of the 809-cm^{-1} mode and conformation was given by Thomas and Peticolas, (1983a). The frequency of this C3′-endo band was correlated with the sugar torsion angle δ of RNA oligomers. Raman studies of B-type DNA crystal structures and normal mode analysis should illuminate this issue.

Arnott et al. (1983) proposed a heteronomous secondary structure for both high- and low-humidity fibers of poly[d(A)]·poly[d(T)]. One strand, probably poly[d(A)], had

C3′-endo sugar puckers and backbone torsion angles similar to those of A-type DNA. The poly[d(T)] strand had C2′-endo sugar puckers and a backbone similar to the canonical B-type conformation. If this conformation occurs in solution, the fraction of deoxyribose rings with an A-like pucker would be expected to be greater than is observed for other DNAs. Although the frequency of the 816-cm^{-1} band of poly[d(A)]·poly[d(T)] implies unique characteristics for this DNA's backbone conformation, the intensity of this band is similar to that of the C3′-endo family bands of all other DNAs (Figure 8). The validity of relating the intensity of this band to the fraction of C3′-endo sugars is indicated by a comparison of A-type DNA fibers and duplex RNAs. These nucleic acids have C3′-endo family bands with nearly the same intensity relative to the 1099-cm^{-1} band (1.6–1.65), even though their peaks differ by 6–8 cm^{-1} (Prescott et al., 1984; Thomas & Hartman, 1973). If poly[d(A)]·poly[d(T)] was 50% C3′-endo, one could expect the relative intensity at 816 cm^{-1} to be about 0.8 instead of 0.32. The width of the 816-cm^{-1} band is narrower than that of the 809-cm^{-1}-type bands of natural DNAs (20 vs. 25–33 cm^{-1}), but it is similar to the widths of other synthetic DNAs.

Thomas and Peticolas (1983a) have shown that the 816-cm^{-1} peak sharpens and increases by ~10% when the temperature of poly[d(A)]·poly[d(T)] is reduced to 0 °C. Jolles et al. (1985) has recently obtained evidence that at 0 °C the adenine strand adopts an A-family structure as suggested by Arnott et al. (1983). Resonance Raman intensities of three adenine modes in poly[d(A)]·poly[d(T)] and other RNA and DNA polymers were compared. The hypochromism of these bands at 0 °C relative to 80 °C implied that the structure of the adenine strand of poly[d(A)]·poly[d(T)] is more like that of poly[r(A)] and poly[r(A)]·poly[r(U)], than poly[d(A)] or poly[d(A-T)]·poly[d(A-T)]. Taken together, the Raman studies indicate that at 0 °C poly[d(A)]·poly[d(T)] has features characteristic of a heteronomous structure. Even at this temperature, however, the relative intensity of the C3′-endo band is less than expected for a rigid heteronomous structure. At 25 °C, the intensity measurements at 816 and 841 cm^{-1} do not indicate a significantly higher ratio of C3′-endo/C2′-endo conformers than for other DNAs. On a Raman time scale, the sugar pucker of poly[d(A)]·poly[d(T)] is more dynamic than indicated by the heteronomous model. NMR studies are also in disagreement with the heteronomous model (Sarma et al., 1985).

In addition to the anomalous frequency of the C3′-endo family band, the intensities of several base vibrations of poly[(dA)·poly[d(T)] differ from the natural DNA extrapolation. The adenine band at 729 cm^{-1} and the thymine bands at 669 and 750 cm^{-1} have slightly higher intensities than the line extrapolated from the natural DNA data (Figures 3 and 4). Base modes at 1302 and 1375 cm^{-1} show lower intensities. Linear dichroism studies of poly[d(A)]·poly[d(T)] in solution indicate that its base pair propeller twist is significantly different from other DNA sequences (Edmondson & Johnson, 1985). The Raman intensity results probably reflect these differences in adenine and thymine modes.

The Raman spectra of poly[d(A-T)]·poly[d(A-T)] has several unique features. It was the only DNA that required three bands in the 800–850-cm^{-1} region. These three bands suggest the presence of three discrete backbone bands. The absence of the "extra" 823-cm^{-1} band in poly[d(A)]·poly[d(T)] argues against it being a base vibration. Although a novel proposal, previous studies indicate some uncertainty as to the backbone of poly[d(A-T)]·poly[d(A-T)]. NMR studies by Assa-Munt and Kearns (1984) and Jenkins et al. (1985) showed that the solution structure of poly[d(A-T)]·poly[d(A-T)] was consistent with a right-handed B-type conformation in which both adenine and thymine have C2~-endo-type sugar puckers. In comparing the NMR studies with our Raman data, it is important to note that these two methods measure signals on different time scales. A Raman scattering event takes place in about 10^{-12}–10^{-13} s, while NMR measurements using the two-dimensional nuclear Overhauser effect average over at least 10^{-6}. If, for example, the deoxyribose ring pucker is fluctuating between conformational states at a rate r, with $10^7 < r < 10^{11}$ s^{-1}, NMR measurements will see an average over two (or more) states whereas Raman spectroscopy can delineate populations of discrete states. The occurrence of three Raman backbone modes is thus not necessarily in conflict with the NMR data. Assa-Munt and Kearns (1984) observed an interaction between two different adenine H2 resonances. This implied a cross-strand proximity of two adenines in different environments and suggested different conformation coexist within poly[d(A-T)]·poly[d(A-T)]. Other evidence for multiple backbone conformations in this polymer comes from the X-ray diffraction studies of Millane et al. (1984). At low relative humidity, a fiber of poly[d(A-T)]·poly[d(A-T)] was observed to have a structural repeat of six nucleotide pairs.

The intensities of base vibrations in poly[d(A-T)]·poly[d(A-T)] also indicate unique features. The 1342- and 1512-cm^{-1} adenine bands have intensities that are greater than those of poly[d(A)]·poly[d(T)] and the line extrapolated from the natural DNAs. A lack of knowledge about the detailed atomic motions of these adenine modes precludes further interpretation. We note though that these two adenine modes show no intensity changes upon DNA denaturation, whereas other adenine bands at 1375 and 1512 cm^{-1} do (Erfurth & Peticolas, 1975). Benevides and Thomas (1985) have also shown that the adenine base in poly[d(A-T)]·poly[d(A-T)] has anomalous properties with respect to deuterium exchange.

The remaining two synthetic DNAs examined in this study, poly[d(T-G)]·poly[d(C-A)] and poly[d(G-C)]·poly[d(G-C)], are clearly of the B type. Most Raman bands have frequencies and intensities consistent with the natural DNA data. A few exceptions are worth noting. Poly[d(G-C)]·poly[d(G-C)] has a higher intensity at 835 cm^{-1} than extrapolated from the natural DNAs (Figures 8). This may indicate an increased population of C2′-endo furanose rings due to the increased stacking interaction in this DNA. A pTpT crystal with well-defined C2′-endo sugar puckers has a peak at 835 cm^{-1} with an intensity much higher than that found for B-DNA fibers (Thomas & Peticolas, 1983a). This observation indicates that this band's intensity is sensitive to the population of deoxyribose rings with a C2′-endo pucker. Stacking may also account for the decrease in the 782-cm^{-1} cytosine band. For poly[d(T-G)]·poly[d(C-A)], the bands that have significant deviations from the expected intensities are adenine and thymine bands (669, 750, 1302, and 1342 cm^{-1}). This implies that the A·T pair of this polymer differs in some way from an average A·T in natural DNAs.

Added in Proof

Recently, G. J. Thomas and collaborators have also obtained evidence for multiple Raman bands in the 800–850-cm^{-1} region for DNA in solution (Thomas et al., 1986).

Acknowledgments

We acknowledge Dr. Ian Gatland for assistance in the initial stages of development of the analysis program. We thank Dr. Bruce Jenkins and Dr. James Alderfer for their generous gift

of poly[d(T-G)]·poly[d(C-A)].

Supplementary Material Available

Figures 1S–4S showing the background curve and 18 bands that fit the 640–1150-cm^{-1} region, the fitted spectra for five and six bands in the *C. perfringens* DNA 700–876-cm^{-1} region, and intensity vs. percent GC plots for eight bands from 1050 to 1530 cm^{-1} (6 pages). Ordering information is given on any current masthead page.

Registry No. Poly[d(A)]·poly[d(T)], 24939-09-1; poly[d(T-G)]·poly[d(C-A)], 27732-52-1; poly[d(A-T)]·poly[d(A-T)], 26966-61-0; poly[d(G-C)]·poly[d(G-C)], 36786-90-0.

References

Arnott, S., Chandrasekaran, R., Hall, I. H., & Puigjaner, L. C. (1983) *Nucleic Acids Res. 11*, 4141.

Assa-Munt, N., & Kearns, D. R. (1984) *Biochemistry 23*, 791.

Benevides, J. M., & Thomas, G. J., Jr. (1985) *Biopolymers 24*, 667.

Chou, C. H., & Thomas, G. J., Jr. (1977) *Biopolymers 16*, 765.

Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Frateni, A. V., & Kopka, M. L. (1982) *Science (Washington, D.C.) 216*, 475.

Drew, H. R., Wing, R. M., Takana, T., Broka, T., Tanaka, S., Itakura, K., & Dickerson, R. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2179.

Edmondson, S. P., & Johnson, W. C., Jr. (1985) *Biopolymers 24*, 825.

Englander, S. W., & Kallenbach, N. (1983) *Q. Rev. Biophys. 16*, 521.

Erfurth, S. C., & Peticolas, W. L. (1975) *Biopolymers 14*, 247.

Erfurth, S. C., Bond, P. J., & Peticolas, W. L. (1975) *Biopolymers 14*, 1245.

Fidor, S. P., Starr, P. A., & Spiro, T. G. (1984) *Biophys. J. 45*, 119a.

Forrest, G., & Lord, R. C. (1977) *J. Raman Spectrosc. 6*, 32.

Hagerman, P. (1981) *Biopolymers 20*, 1503.

Hartman, K. A., Lord, R. C., & Thomas, G. J., Jr. (1973) in *Physico-Chemical Properties of Nucleic Acids* (Duchesne, J., Ed.) pp 1–89, Academic, New York.

Horowitz, D. S., & Wang, J. C. (1984) *J. Mol. Biol. 173*, 75.

Jenkins, B., & Alderfer, J. L. (1985) *Biophys. J. 47*, 331a.

Jolles, B., Laigle, A., Chinsky, L, & Turpin, P. Y. (1985) *Nucleic Acids Res. 13*, 2075.

Klysik, J., Stirdivant, S., Larson, J. E., Hart, P. A., & Wells, R. D. (1981) *Nature (London) 290*, 671.

Lord, R. C., & Thomas, G. J., Jr. (1967) *Spectrochim. Acta, Part A 23A*, 2551.

Lu, K.-C., Prohofsky, E. W., & Van Zandt, L. L. (1977) *Biopolymers 16*, 2491.

Lu, K.-C., Van Zandt, L. L., & Prohofsky, E. W. (1979) *Biophys. J. 28*, 27.

Marquardt, J. W. (1963) *J. Soc. Ind. Appl. Math. 11*, 431.

Martin, J. C., & Wartell, R. M. (1982) *Biopolymers 21*, 499.

Milane, R. P., Walker, J. K., Arnott, S., Chandrasekaran, R., & Birdsall, D. L. (1984) *Nucleic Acids Res. 12*, 5475.

Mirau, P. A., & Kearns, D. R. (1983) in *Structure and Dynamics: Nucleic Acids and Proteins* (Clementi, E., & Sarma, R. H., Eds.) pp 227–239, Adenine, Albany, NY.

Mirau, P. A., & Kearns, D. R. (1985) *Biopolymers 24*, 711.

Nishimura, Y., Hirakawa, A. Y., & Tsuboi, M. (1978) *Adv. Infrared Raman Spectrosc. 5*, 217.

Nishimura, Y., Tsuboi, M., & Sato, T. (1984) *Nucleic Acids Res. 12*, 6901.

Pohl, F. M., Ranade, A., & Stockburger, M. (1973) *Biochim. Biophys. Acta 335*, 85.

Prescott, B., Steinmetz, W., & Thomas, G. J., Jr. (1984) *Biopolymers 23*, 235.

Reid, D. G., Salisbury, S. A., Bellard, S., Shakked, Z., & Williams, D. H. (1983) *Biochemistry 22*, 2019.

Rich, A., Quigley, G. J., & Wang, A. H.-J. (1981) in *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) pp 35–52, Adenine, Albany, NY.

Sarma, M. H., Gupta, G., & Sarma, R. H. (1985) *J. Biomol. Struct. Dyn. 2*, 1057.

Shakked, D., Rabinovich, D., Kennard, O., Cruse, W. B. T., Salisbury, S. A., & Viswamitra, M. A. (1983) *J. Mol. Biol. 166*, 183.

Shore, D., & Baldwin, R. L. (1983) *J. Mol. Biol. 170*, 957.

Thomas, G. J., Jr., & Hartman, K. C. (1973) *Biochim. Biophys. Acta 312*, 311.

Thomas, G. J., Jr., & Kyogoku, Y. (1977) *Pract. Spectrosc. 1C*, 717.

Thomas, G. A., & Peticolas, W. L. (1983a) *J. Am. Chem. Soc. 105*, 986.

Thomas, G. A., & Peticolas, W. L. (1983b) *J. Am. Chem. Soc. 105*, 993.

Thomas, G. J., Jr., & Benevides, J. M. (1985) *Biopolymers 24*, 1101.

Thomas, G. J., Jr., Benevides, J. M., & Prescott, B. (1986) in *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) Vol. III, Adenine, Albany, NY.

Strommen, D. P., & Peticolas, W. L. (1982) *Biopolymers 21*, 969.

Wang, A. H.-J., Quigley, G. J., Kolpak, G., van der Marel, J. H., van Boom, J. H., & Rich, A. (1980) *Science (Washington, D.C.) 211*, 171.

Wartell, R. M., & Benight, A. S. (1982) *Biopolymers 21*, 2069.

Wartell, R. M., Harrell, J. T., Zacharias, W., & Wells, R. D. (1983) *J. Biomol. Struct. Dyn. 1*, 83.

THERMAL DENATURATION OF DNA MOLECULES: A COMPARISON OF THEORY WITH EXPERIMENT

Roger M. WARTELL and Albert S. BENIGHT

School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A.

NORTH-HOLLAND–AMSTERDAM

PHYSICS REPORTS (Review Section of Physics Letters) 126, No. 2 (1985) 67–107. North-Holland, Amsterdam

THERMAL DENATURATION OF DNA MOLECULES: A COMPARISON OF THEORY WITH EXPERIMENT

Roger M. WARTELL and Albert S. BENIGHT

School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A.

Received February 1985

Contents:

Abstract:

Experimental and theoretical results on the helix-coil transition of DNA are reviewed. The theoretical model of the transition is described, and the influence of heterogeneous base pair stacking, and strand dissociation on the predicted melting transition is examined. New experimental transition data on seven DNAs, 154–587 base pairs (bp) long, are reported and compared with theoretical calculations. We review and evaluate previous studies on long DNAs (≥1000 bp) as well as previous and recent results on short DNAs. The comparison of theory with equilibrium melting curves of short DNAs indicates that base pair sequence has a relatively small influence on the stacking free energy. Excellent agreement is obtained between theory and equilibrium transitions of 14 out of 15 fragments 80–587 pb. The deviation between theory and experiment for a 516 bp DNA can be attributed to the formation of stem-loop structures. This may provide the explanation for inconsistent results observed with long DNAs. The effect of single base pair changes on DNA transitions is discussed. Current views on fluctuational opening of base pairs at temperatures below the transition are described.

Single orders for this issue

PHYSICS REPORTS (Review Section of Physics Letters) 126, No. 2 (1985) 67–107.

Copies of this issue may be obtained at the price given below. All orders should be sent directly to the Publisher. Orders must be accompanied by check.

Single issue price Dfl. 26.00, postage included.

1. Introductio

When DN
structure is a
meandering v
together by i
biological pro
DNA into sin
time to every
base pairs, or

An analogu
middle 1950s s
in a cooperati
one-half of the
between 50°C
physical under
the biological
more complex
the helix-coil
insights on the
unwinding, an
understanding
since the late
melting of DN
tractable mode
results have co

This report
DNA. New exp
theoretical calc
is placed on e
which occur be
pair changes o
DNA gel electr
migrating DNA
one base pair.
to develop a
fulfilled in an
temperatures b
binding protein
several years o
reviews for add
1970s [11–13].

1.1. Backgroun

A DNA mel

. Introduction

When DNA is placed in solution under physiological solvent conditions, its average secondary tructure is a two stranded helix. Superficially one can describe a DNA molecule in solution as a neandering worm-like entity comprised of two tightly intertwined strands. The strands are held ogether by interactions between complementary pairs of base units (fig. 1). If one considers the biological processes involving DNA, e.g., the duplication of DNA, or the transcription or copying of DNA into single stranded RNA, it is clear that in a cell localized unwinding occurs at some point in ime to every DNA region. This unwinding may take the form of fluctuational opening of individual base pairs, or the strand separation of large regions driven by DNA binding proteins.

An analogue of DNA unwinding in a cell can be observed with solutions of purified DNA. In the middle 1950s several workers recognized that heating DNA solutions above room temperature resulted in a cooperative unraveling of the intertwined strands to single strands [1, 2]. The temperature at which one-half of the DNA denatured, T_m, depended on the specific DNA and solvent. The T_m generally fell between 50°C and 100°C. Since these initial studies were made, a vast literature has contributed to a physical understanding of DNA unwinding. Although the mechanism of unwinding clearly differs from the biological protein mediated process, studies on DNA alone are a first step toward understanding the more complex situation. The cooperative unwinding of two stranded DNA to single strands is known as the helix-coil or melting transition. Investigations of this transition have provided information and insights on the interactions governing DNA unwinding, the influence of base pair sequence on regional unwinding, and the influence of the solvent on DNA stability [3, 4, 14, 42]. A semiquantitative understanding of DNA melting, based on the Ising model of statistical physics [5, 6] has been available since the late 1960s. A detailed quantitative understanding of the transition has remained elusive. The melting of DNAs of different lengths and base pair sequences present theorists with a challenge. Can a tractable model be produced which accurately predicts observed transition curves? Although new results have come teasingly close, this challenge has not yet been fully met.

This report will review recent experimental and theoretical studies of the helix-coil transition of DNA. New experimental transition data on a number of short DNAs are described and compared with theoretical calculations. Our findings as well as those of other workers are critically evaluated. Emphasis is placed on experiments done under equilibrium conditions. Possible reasons for the discrepancies which occur between theory and experiment are discussed. We also examine the influence of single base pair changes on the melting behavior of DNA. This analysis has relevance to recent developments in DNA gel electrophoresis [7, 8]. Lerman has employed chemical denaturants in a gel to induce melting in migrating DNA molecules. This process can separate 536 bp long DNAs which differ in sequence by one base pair. The realization of one of the original motivations for studying the helix-coil transition – to develop a method for determining base sequence information from transition curves – may be fulfilled in an unexpected way. The final area of this report will focus on base pair opening at temperatures below the transition region. This phenomenon is important to understand what a DNA binding protein 'sees' when it approaches a DNA site. Two other reviews have appeared in the last several years on the DNA helix-coil transition [9, 10]. The interested reader is urged to consult these reviews for additional perspectives on this topic. Earlier work in the field has been reviewed in the early 1970s [11–13].

1.1. Background

A DNA melting curve is generally a two-dimensional plot displaying some property of a DNA

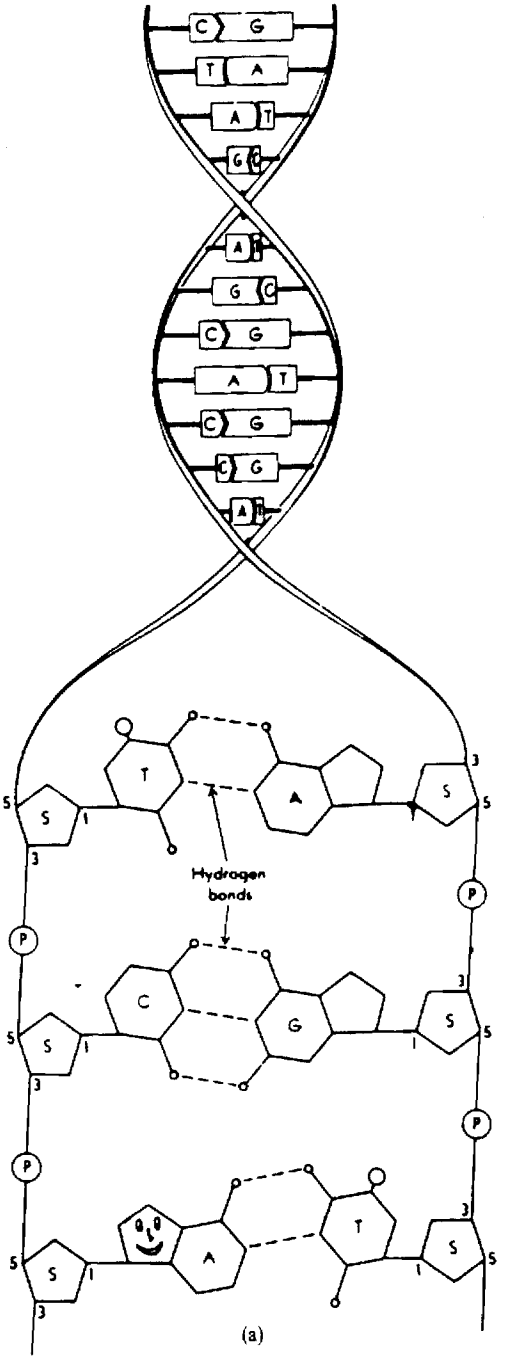


Fig. 1. Two schematic representations of a portion of a DNA duplex. (a) The hydrogen bonds connecting the base units adenine, A, and thymine, T, or guanine, G, and cytosine, C, are shown as dashed lines. S represents the deoxyribose sugar group and P the phosphate group. Base stacking interactions also stabilize the double helix. (b) Space filling molecular model picture of 'B'-conformation DNA.

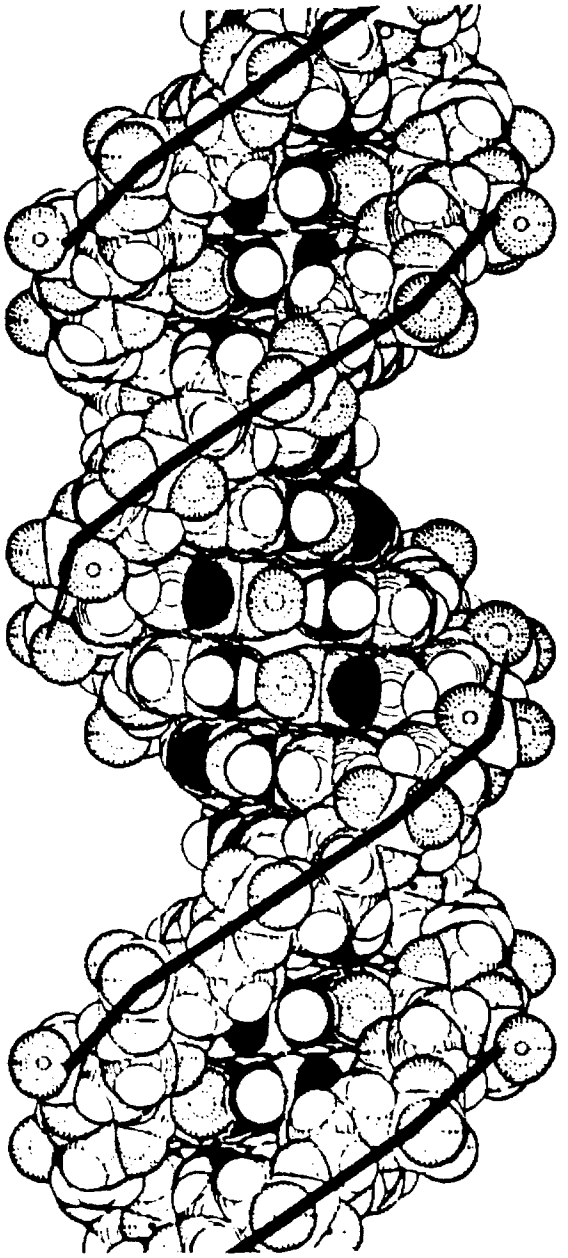


Fig. 1(b)

solution against an external variable producing DNA unwinding. The expernal variable is most commonly the temperature, but 'melting' can also be brought about by extremes of pH, decrease in dielectric constant of the aqueous medium by alcohol, ketones, etc., exposure to amides, urea and similar solvents [8, 9].

The most popular method of investigating the melting transition in DNA has been to observe the change in UV absorbance of a dilute DNA solution while slowly elevating the temperature. By monitoring the optical absorbance at approximately 260 nm while increasing the temperature, and normalizing the absorbance change in an appropriate way, a plot of θ_B, the fraction of broken base pairs, versus temperature can be calculated. Figure 2 illustrates a helix-coil transition which might be observed for a short DNA (≤ 200 bp). In the early part of the transition there is a slight linear increase in absorbance. This linear increase is interpreted as a slight linear increase in the average base pair stacking of the double helix. The DNA is still predominantly base paired. Absorbance values in this pre-transition linear region are specified by a linear function of temperature, $A_L(T)$.

As the temperature is raised, hydrogen bonded base pairs between complementary strands begin to open inducing the strands to unwind. Opening of the remaining base pairs occurs in a highly cooperative manner causing the strands to completely separate. After strand separation, a linear absorbance is again observed if the temperature is further increased. This absorbance corresponds to the increased unstacking of bases in the single strands and is represented by $A_U(T)$. The fraction of broken base pairs θ_B vs. temperature is obtained by correcting the observed absorbance change for volume expansion and plotting

$$\theta_B = [A(T) - A_L(T)]/[A_U(T) - A_L(T)]\,. \tag{1}$$

Melting curves are also commonly expressed in terms of the fraction of intact base pairs, $\theta = 1 - \theta_B$. Where one should draw $A_L(T)$ is not always apparent, and some arbitrariness exists regarding the beginning of DNA melting. Fortunately this uncertainty does not have a large effect on the normalized melting curve for DNAs over 100 bp.

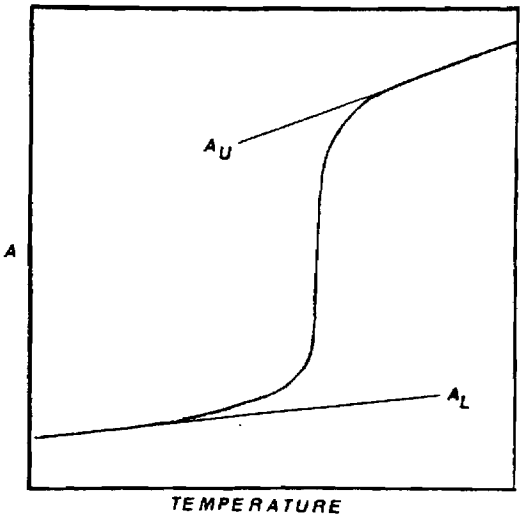


Fig. 2. A typical plot of absorbance vs. temperature demonstrating the denaturation of a short DNA. Absorbance increase is normally 30–40% from a solution with 0.4 absorbance units at 260 nm.

The UV absorbance of DNA is due to the nucleotide bases. It arises from the π–π^* electronic transition in both purine and pyrimidine bases [14]. An increase in the absorbance reflects a change in the electronic configuration of the bases due to the decrease in double helical stacking. The absorbance increases by 30–40% depending on the DNA sample. As implied by eq. (1) it is generally assumed that a linear relationship exists between increased absorbance and the extent of disrupted base pairs. This appears to be true if the disrupted helical region contains at least ten base pairs [14]. The percentage increase in light absorption at 260 nm produced by denaturing DNA is related to the DNA's percentage of A · T base pairs. The higher the percentage of A · T's in a DNA, the greater the percentage increase of absorbed light [15]. If the absorption is monitored at 268 nm, the percentage increase in hyperchromicity for unwinding A · T or G · C base pairs is approximately the same [15]. Thus 268 nm is the optimum wavelength for monitoring a DNA melting curve to obtain θ_B.

In the 1960s, studies on DNA melting were carried out with sheared chromosomal and virial DNAs from various sources. One of the important findings of that period was the correlation observed between the midpoint temperature of the transition of a DNA, T_m, and its % G · C content [16, 17]. The T_m's of natural DNAs increased linearly with increasing % G · C. This indicated that G · C base pairs were, on the average, more stable than A · T pairs. Another observation noted was that distinct multistep melting curves occurred for a few DNAs [18, 19]. Although most DNA samples gave smooth sigmoidal transitions due to their heterogeniety in length and base pair sequence, DNAs from viruses showed multistep transitions. Developments in DNA isolation procedures in the 1970s made it possible to study a number of DNA samples which were homogeneous in length and sequence. Multistep melting curves became the normal observation [10]. This multistep behavior is best manifested by plotting the temperature derivative of the fraction of broken base pairs vs. temperature. Figure 3 shows an example of this type of data. A plot of absorbance vs. temperature and the corresponding differential melting curve for a 1630 bp DNA are presented. Seven distinct peaks are observed. An estimate of the

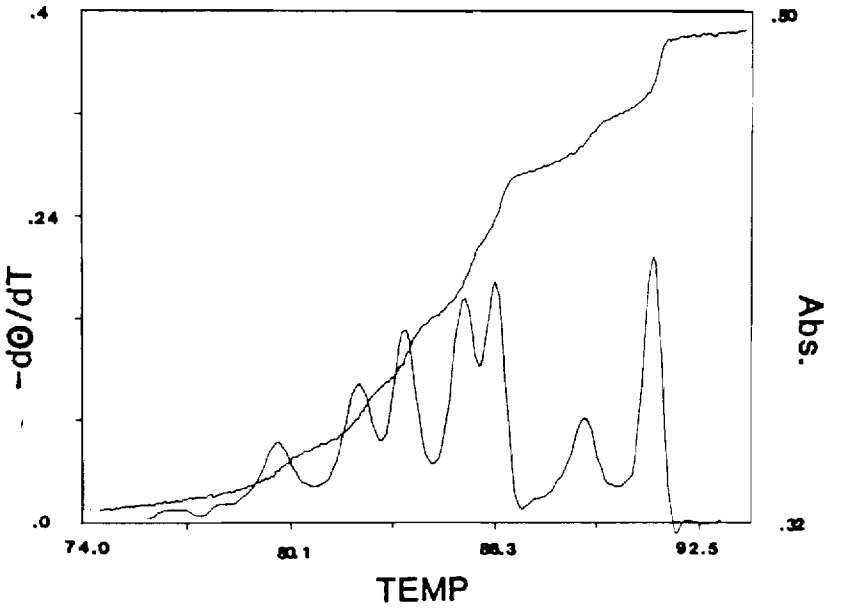


Fig. 3. Absorbance vs. temperature and differential melting curve for the 1630 bp *Hinf* I restriction endonuclease DNA fragment of the plasmid pBR322 (85).

number of base pairs melting in each step of the transition may be obtained by deconvoluting the differential melting curve into a superposition of seven Lorentzian curves. Normalizing the areas of these curves to the 1630 bp length, one finds that the cooperatively melting regions correspond to 100–350 bp in length. Cooperatively melting regions have also been called thermalites. One of the interesting characteristics of biological macromolecules is the interplay between the cooperative interactions between regions and the independent properties of these regions. The transition shown in fig. 3 exemplifies this for DNA. Regions of DNA on the order of one hundred base pairs unwind over temperature intervals of 0.3–0.7°C. When a specific region opens the stability of adjacent regions is decreased. The progress of unwinding depends on the cooperativity between regions and the internal stability of each region. Electron microscopy has been used to visualize and verify the stepwise unwinding of DNA regions [21, 22].

In order to establish a quantitative understanding of the processes of DNA melting, it is necessary to test theoretical predictions against experimental curves such as fig. 3. The basic theory of the DNA helix-coil transition was developed in the 1960s and early 1970s. Meaningful comparison with experimental curves of natural DNAs had to wait until the development of techniques to determine the complete base pair sequence of DNAs. This enabled workers to obtain differential melting curves (DMC) from DNAs of known sequence and to compare them to theoretical curves. References [9] and [10] have reviewed results on DNAs from fd, $\phi \times 174$ and T7 viruses as well as several plasmid DNAs. The lengths of these DNAs range from 2 to 6 kilobase pairs (kbp). Before discussing the theory–experiment comparison of these DNAs as well as those of shorter DNAs (80–567 bp) we will review the theoretical model of DNA melting and examine several variants of the basic formulation. Although the theory has been reviewed by several authors we wish to provide a common background to the uninitiated reader, and focus on several points which are relevant to later discussions.

2. Theoretical calculation of DNA melting

All theoretical approaches investigating the helix-coil transition of DNA employ the same fundamental model. DNA is considered as a quasi-one-dimensional lattice composed of N base pair units. The model recognizes that the double helix is held together by hydrogen bonds between complementary bases on opposite strands (A · T or G · C) and hydrophobic or stacking interactions between nearest neighbor bases on opposite strands. Each base pair is in one of two states: hydrogen bonded (intact) or non-hydrogen bonded (open). The molecular structure correlated with the open state is not known. Empirically the open state can be regarded as the structure which generates an absorbance increase at 268 nm attributable to one base pair. One must have the bases sufficiently unstacked from neighboring bases to generate an increased absorbance in the $\pi \rightarrow \pi^*$ electronic transition. This point will be considered again later with regard to fluctuational base pair opening below the melting region.

Characterizing the DNA molecule in a two-state manner is similar to the characterization employed in the well-known Ising model of ferromagnetism [5]. In addition to the self-energy term of each base pair and the nearest neighbor interaction, however, one must also include two additional features to describe a DNA transition. One is the configuration entropy difference between two unbonded strands at the end of a DNA and an equivalent sized unbonded region or loop sandwiched between bonded segments. The second is the dissociation equilibrium which results in the separation of the two strands when the last base pair(s) open.

The average fraction of intact base pairs, $\theta(T)$ for an ensemble of DNA molecules of identical

sequence is given by

$$\theta(T) = \theta_{\text{int}}\theta_{\text{ext}} . \tag{2}$$

θ_{int} is the average fraction of intact base pairs among those DNAs with at least one base pair. These duplex DNAs will be represented by C_2. θ_{ext} is the fraction of DNA *strands* which have at least one intact base pair with a complementary strand. Single strands will be represented by C_1. θ_{ext} is related to the dissociation equilibrium for duplex DNA to single strands $C_2 \rightleftarrows 2C_1$. This dissociation step is generally ignored for the melting of long DNAs since θ_{ext} is close to 1.0 until θ_{int} is nearly zero. For short DNAs ($\leq$200 bp) fluctuations about the equilibrium thermal energy can cause a significant fraction of DNA molecules to dissociate to single strands even with θ_{int} around 0.5 [23]. Inclusion of the dissociation step is essential to accurately describe the melting behavior of short DNA molecules. The calculation of θ_{int} will be treated first. It depends on a model describing changes in the internal degrees of freedom of a DNA with at least one base pair. The internal partition function is denoted as $Z(C_2)$.

2.1. DNA melting prior to dissociation

For any given temperature, one must determine the probability of each configuration available to the DNA duplex. Each DNA configuration k will have N_k intact base pairs. The average fraction of intact base pairs is given by

$$\theta_{\text{int}} = \sum_{\{k\}} [N_k Z_k / Z(C_2)]/N, \tag{3}$$

where Z_k is the statistical weight for configuration k and $\{k\}$ is the sum over all configurations.

2.1.1. Parameters in heterogeneous stacking model

Three types of parameters are employed to describe the statistical weights of the configurations accessable to an N base pair DNA. For a given configuration k, Z_k is a product of stability parameters s – one for each intact base pair, cooperativity parameters σ – one for each internal loop of unbonded base pairs, and loop entropy parameters $f(m)$, one for each internal loop of m unbonded base pairs. Figure 4 shows one configuration for a six base pair DNA and its statistical weight. Open base pairs at the ends of a DNA are given the statistical weight one. The stability parameter s is actually an equilibrium constant assigned to the reaction of forming a stacked hydrogen bonded base pair at one end of a duplex segment. Within the framework of the nearest neighbor model the net free energy change of this process depends on the base pair being formed (A·T or G·C), and the stacking interaction formed. Thus for the reaction of closing the ith base pair onto the $(i-1)$ base pair (fig. 5)

$$s_i' = \exp[-(G_i + G_{i-1,i}^{s})]/RT, \tag{4}$$

G_i is the free-energy difference between the intact state and the open state attributable to only the ith base pair. This includes interactions such as solvent effects as well as base pair hydrogen bonding. $G_{i-1,i}^{s}$ includes the stacking free-energy difference between intact and open states, and the entropy contribution for positioning the bases for pairing. We note that closing the ith base pair can generate a different s value

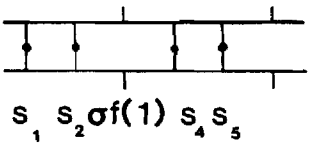


Fig. 4. A schematic representation of a six base pair segment and the statistical weight of the configuration shown.

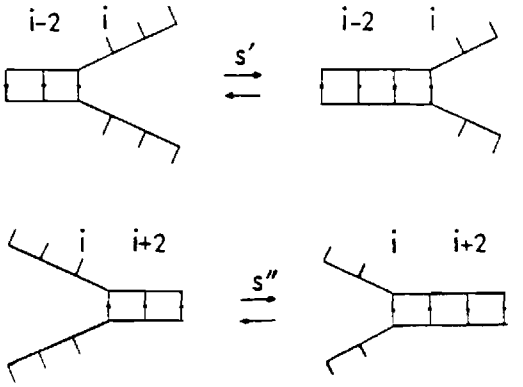


Fig. 5. Two reactions for making an intact ith base pair. See text for s' and s''.

depending on the stacking interaction, i.e., s' and s'' in fig. 5. A cooperativity parameter σ and loop entropy parameter $f(m)$ occur whenever there is an internal loop of unbonded base pairs. The σ's correspond to the free energy changes $G^s_{i-1,i}$ that occur in the reactions described by s,

$$\sigma_{i-1,i} = \exp(G^s_{i-1,i}/RT)\,. \tag{5}$$

They account for the fact that two stacking interactions are broken when an internal base pair opens whereas only one stacking interaction is lost when an end base pair opens. $f(m)$ is defined as the probability that two complementary strands joined at a duplex are in a spatial arrangement that allows the formation of one bonded base pair after a loop with m unpaired bases in each strand. The loop closure parameter depends on the 'stiffness' of the DNA strands or more formally their persistence length. Slightly different definitions of σ and $f(m)$ are used by others. This is discussed in ref. [24].

The simplest assumption to make with regard to base pair stability is to consider the stacking interaction, $G^s_{i,i+1}$, to be the same for all base pairs. Thus from eqs. (4) and (5) one has a single σ value and two different s values; s_{AT} for an A·T base pair, and s_{GC} for a G·C base pair. This assumption with an appropriate $f(m)$ has been employed in many calculations of DNA melting curves [20, 23, 25, 26]. The observed linear correlation of natural DNA T_m's with their G·C composition suggests this assumption is valid at least to first order.

Gotoh and Tagashira employed a model in which the stability parameters s are assigned to base pair doublets rather than base pairs [9, 27]. This calculation assumed that the G_i's in eq. (4) are the same for A·T and G·C pairs and differences in base pair stability are due to stacking interactions. This assumption is also made in the analyses of oligonucleotide melting curves [28, 29]. Although experimental [30] and theoretical [31] studies indicate that stacking interactions differ for different base pair doublets, no evidence indicates that the hydrogen bond free energy change in opening A·T and G·C pairs is the same. On the contrary, theoretical studies [32] show that the hydrogen bond energy values are significantly different for G·C vs. A·T base pairing. We will employ a model which includes the sequence dependence of both the hydrogen bond energy terms and the stacking energy terms. Since there are 10 different nearest neighbor pairs, $G^s_{i,i+1}$ will take on one of ten values. G_i has one of two possible values.

The Ising model formulation familiar to physicists would designate a configuration of an N base pair

DNA by a sequence of τ_i's,

$$(\tau_1, \tau_2, \ldots, \tau_i, \ldots, \tau_N), \tag{6}$$

where

$$\tau_i = \begin{cases} +1 & \text{if the } i\text{th base pair is bonded}, \\ -1 & \text{if the } i\text{th base pair is open}. \end{cases}$$

The equilibrium probability of finding the DNA in a particular configuration k at a given temperature is obtained from the ratio of the statistical weight Z_k and the partition function Z. For the interaction model described above, Z is given by

$$Z = \sum_{\{\tau_i\}} \prod_{i=1}^{N} P(\tau_i, \tau_{i+1}) \prod_{\{m_j\}} f(m_j), \tag{7}$$

where the set $\{\tau_i\}$ describes all configurations with at least one base pair and the set $\{m_j\}$ is the set of loops corresponding to a specific configuration $(\tau_1, \tau_2, \ldots, \tau_N)$. $P(\tau_i, \tau_{i+1})$ the probability function for (τ_i, τ_{i+1}) is given by

$$P(\tau_i, \tau_{i+1}) = \exp\left[\frac{(1+\tau_i)(1+\tau_{i+1})G^s_{i,i+1}}{4RT} + \frac{(1+\tau_i)G_i + (1+\tau_{i+1})G_{i+1}}{4RT}\right] \tag{8}$$

for $2 \le i \le N-1$. For the base pairs at the ends one stacking interaction is missing.

In order to construct an algorithm which generates a unique statistical weight for each DNA configuration, one must assign a single stability parameter s_i for each base pair. This parameter is the average of the two reactions shown in fig. 5, s'_i and s''_i, or

$$s_i = \exp - [G_i + \tfrac{1}{2}(G^s_{i,i-1} + G^s_{i,i+1})]/RT. \tag{9}$$

G_i has one of two possible values, depending on the ith base pair type, A · T or G · C. $G^s_{i,i-1}$ has one of ten possible values. If we write $G^s_{i,i-1} = \overline{G^s} + \delta G_{i,i-1}$ where $\overline{G^s}$ is the average of the ten stacking energy terms then

$$s_i = \exp[-\{(G_i + \overline{G^s}) + \tfrac{1}{2}(\delta G_{i,i-1} + \delta G_{i,i+1})\}/RT]. \tag{10}$$

The term $(G_i + \overline{G^s})$ in eq. (10) corresponds to the free energy change of average A · T or G · C pairs in a random sequence DNA, ΔG_{AT} or ΔG_{GC}. The δG terms provide the nearest neighbor sequence dependence on s_i.

Proper consideration of heterogeneous stacking is obtained by assigning a specific cooperativity parameter $\sigma^{1/2}_{i,i+1}$ at each junction between an intact and opened base pair. This cooperativity parameter is given in terms of $\overline{G^s}$ and δG

$$\sigma^{1/2}_{i,i+1} = \tilde{\sigma}^{1/2} \exp(\delta G_{i,i+1}/2RT) = \exp(\overline{G^s}/2RT) \exp(\delta G_{i,i+1}/2RT). \tag{11}$$

The inclusion of sequence specific $\sigma^{1/2}$ values enables one to assign correct equilibrium constants for opening a given base pair from both its left and right side.

2.1.2. Relation of stability parameters to enthalpy and entropy

The averaged equilibrium constant s_i of eq. (10) can be expressed as

$$s_i = \exp[-(\Delta H_i - T\,\Delta S_i)/RT]\,, \tag{12}$$

with ΔH_i the enthalpy change and ΔS_i the entropy change averaged over the two reactions of fig. 5. Calorimetric studies on DNA oligomers are determining ΔH_i and ΔS_i for different sequences [29, 33]. Since a complete set is not yet available, eq. (12) will be rewritten in order to relate ΔH_i and ΔS_i to available data. Comparing eqs. (10) and (12) suggests that ΔH_i and ΔS_i be written in terms of the enthalpy and entropy changes of average A · T and G · C pairs and deviations from these averages.

$$\Delta H_i - T\Delta S_i = \begin{cases} \Delta H_{AT} - T\,\Delta S_{AT} + \delta H_i - T\,\delta S_i \\ \Delta H_{GC} - T\,\Delta S_{GC} + \delta H_i - T\,\delta S_i \end{cases} \quad \text{for} \begin{cases} i = \text{AT} \\ \text{GC} \end{cases} \tag{13}$$

$(\delta H_i - T\,\delta S_i)$ is related to the nearest neighbor terms $\delta G_{i,i\pm1}$, and $(\Delta H_{AT} - T\,\Delta S_{AT})$ and $(\Delta H_{GC} - T\,\Delta S_{GC})$ correspond to ΔG_{AT} and ΔG_{GC}, respectively.

Calorimetric data by Klump and Ackerman [34] provide estimates of ΔH_{AT} and ΔH_{GC}, and indicate that $\Delta S_{AT} = \Delta S_{GC} = \Delta S$. The average melting temperatures of an A · T pair, T_{AT}, and a G · C pair, T_{GC}, are related to these thermodynamic parameters by $T_{AT} = \Delta H_{AT}/\Delta S$ and $T_{GC} = \Delta H_{GC}/\Delta S$. For a solvent with 0.1 M Na^+, $T_{AT} = 64.3$°C and $T_{GC} = 107$°C. If the $\delta G_{i,i\pm1}$ terms of eq. (10) are assumed to be independent of temperature, they alter only the enthalpy (i.e., $\delta S_i = 0$ all i). One obtains from eqs. (12) and (13)

$$s_i = \exp\left[\frac{\Delta S}{RT}(T - T_i)\right], \tag{14a}$$

with

$$T_i = \begin{cases} T_{AT} \\ T_{GC} \end{cases} + (\delta G_{i,i-1} + \delta G_{i,i+1})/2\,\Delta S\,. \tag{14b}$$

Several efforts have been made to evaluate the ten δG values, or equivalently the sequence dependence of T_i. Belentsev et al. [35] were the first to estimate these values from periodic base sequence DNA polymer melting curves. The total free energy of a DNA polymer, ΔG_T, is zero at its melting temperature.

If the loop entropy free energy term is neglected, then

$$\Delta G_T(T_m) = N_{AT}\,\Delta G_{AT} + N_{GC}\,\Delta G_{GC} + \sum_{i=1}^{N-1} \delta G_{i,i+1} = 0\,, \tag{15}$$

with N_{AT}, the number of A · T pairs, and N_{GC} the number of G · C pairs. Dividing by the number of base

pairs and expressing ΔG_{AT} and ΔG_{GC} in terms of T_{AT} and T_{GC}, one has

$$\Delta S[(1-f_G)(T_m - T_{AT}) + f_G(T_m - T_{GC})] + \sum_{(MN)} a_{MN}\,\delta G_{MN} = 0\,, \tag{16}$$

where f_G is the fraction of G·C pairs, δG_{MN} is the difference in the stacking free energy of the MN nearest neighbor doublet from G^s, and a_{MN} is the fraction of the MN base pair doublet in the DNA. M and N are the bases in the 5′→3′ direction of one of the two strands, e.g., ApT, TpA, ... In the work by Belentsev et al., certain assumptions were necessary since the DNA polymer data available did not provide an independent set of equations for the 10 different nearest neighbor sequences. Wartell and Benight [24] used the above approach employing additional DNA polymer data together with T_m values of natural DNAs whose nearest neighbor frequencies were known [32, 36]. A least squares fit was made to evaluate δG_{MN}'s from the T_m data of thirteen DNA polymers and four natural DNAs using eq. (16). Table 1 lists two sets of δG parameters determined by this process. Each set obeyed the same convergence criteria. They differ in the initial δG_{MN}'s assumed in the Levenberg–Marquardt algorithm [37]. Set 1 started with $\delta G_{MN} = 0$ for all MN. Set 2 started with the δG_{MN}'s of Gotoh and Tagashira (see table 1). Both sets provided T_m values 2.1°C or less from the 17 experimental T_m's. The standard deviation was about ±0.82°C for both sets.

The large disparity between these two sets indicates that the accuracy of the T_m data and number of T_m values are insufficient to limit the parameter space of the δG_{MN}'s.

Gotoh and Tagashira employed the melting temperatures of 43 cooperatively melting regions in several DNAs to evaluate nearest neighbor interactions [27]. As mentioned earlier, their stability parameter s is assigned to base pair doublets. They assign a melting temperature T_{MN} to the MN doublet. Their T_{MN} can be expressed as $T_{MN} = (T_M + T_N)/2 + \delta T_{MN}$ where T_M and T_N are T_{AT} or T_{GC} depending on whether M or N is an A·T or G·C pair. δT_{MN} is related to δG_{MN} by

$$\delta T_{MN} = \delta G_{MN}/\Delta S\,. \tag{17}$$

Table 1
Stacking interactions
δG_{MN} values (cal/mole) evaluated from a least squares fit of eq. (16) to the T_m's of 18 DNAs are given by sets 1 and 2. Sets 3 and 4 were obtained from the data of refs. (27) and (32), respectively

MN doublet 5′→3′ M_pN	Set 1 (start $\delta G_{MN} = 0$, all MN)	Set 2 (start δG_{MN} = set 3 values)	Set 3 (Gotoh and Tagashira)	Set 4 (Ornstein and Fresco)
A_pA	−54	−54	−80	270.0
A_pT	−112	−246	−143	−1320
A_pG	−23	321	383	−1070.0
A_pC	−33	−510	−608.0	−1330.0
T_pA	267	401	369.0	2200.0
T_pG	−194	283	476	1020.0
T_pC	216	−128	−324.0	+1030.0
G_pG	178	178	249.0	−30.0
G_pC	−271	−1092	−1016.0	−1240
C_pG	151	972	588.0	470

Table 1 lists the δG_{MN} values evaluated from the T_{MN} values of Gotoh and Tagashira. One of the observations made by these authors was that the stacking interactions gave a good prediction of the T_m's of DNA polymers. This is indicated by sets 2 and 3 of table 1. Relatively small changes in the Gotoh and Tagashira parameters provide a least squares minimum to the T_m's of 13 DNA polymer with repeated base pair sequence and four natural DNAs. Another observation was that purine ($5' \rightarrow 3'$) pyrimidine stacking interactions were always more stable than pyrimidine ($5' \rightarrow 3'$) purine stacking. This is observed in their δG_{MN} values, as well as the set 2 values. The binding of drug intercalators to DNA dimer duplexes can be explained by this hierarchy of stabilities [38]. Table 1 also lists one additional set of δG_{MN} values, set 4. These values are the deviations from the average stacking energy calculated by Ornstein and Fresco [32]. They are based on energy minimizations of stacked base pair dimers [31]. Gotoh and Tagashira noted a linear correlation between their T_{MN} values and the energies of nearest neighbor doublets calculated by Ornstein, Rein and coworkers [31]. If, however, one compares the heirarchy of the δG_{MN} values to the recent stacking energy determinations of Ornstein and Fresco [32] a similar correlation is not observed. We note that the latter comparison is more valid since it is for the deviations from the average stacking energy. Base pairing interactions are not considered. Another point worth noting concerns the uniqueness of the δG_{MN} values (or T_{MN}'s) determined by least-squares fitting. Sets 1 and 2 of table 1 show that the initial guess strongly influence the outcome of a non-linear least squares fit. It is not clear if the 43 T_m values employed by Gotoh and Tagashira is a large enough data set to render their result insensitive to the initial guess. Discussion of table 1 parameters with regard to predictions of DNA melting curves is postponed until later.

2.1.3. Loop entropy

The loop entropy term is generally assumed to be proportional to $1/(m+1)^k$ for large loops with m unbonded base pairs with $1.5 \leq k \leq 2.0$. This relationship derives from the Jacobson–Stockmayer calculation for the probability of the ends of a Gaussian polymer of $2m+2$ links meeting within a small volume δv of each other [39]. A more generalized form [40] intended to consider the effects of chain stiffness employs an $f(m)$ of

$$f(m) = 1/(m+d)^k , \tag{18}$$

with d a constant. This formulation was employed by Gotoh and Tagashira [27] to fit experimental melting curves of several DNA molecules (see below). Azbel reexamined the probability of loop closure from a different perspective [41b]. This analysis included the elastic properties of the strands, a physical reality previously omitted. Single strands are considered to be smooth continuous chains of connected segments. The angle between the tangents of segments separated along the strand contour by the distance l, $\theta(l)$, depends on the chain stiffness. For small angles the free energy required to bend the chain by the angle $\theta(l)$ over the distance l is $\Delta G = e_w\theta^2/2l$. e_w is the bending force constant, and bending in only one plane is assumed. The root mean square of $\theta(l)$ is $\langle\theta^2(l)\rangle^{1/2} = (lkT/e_w)^{1/2} = (l/P)^{1/2}$ where P is the persistence length of the strand [14]. Azbel considered the displacement of one of the two strands of a loop from the position it would maintain if it were part of a duplex segment. Calculations were carried out in terms of the root mean square of θ for adjacent segments, called χ, and m the number of bases or segments in one strand of the loop. Results of the analysis give for a loop with m broken base pairs

$$f(m) = \begin{cases} 1 & m \le \chi^{-2/3}, \\ 1/\chi^2 m^3 & \chi^{-2/3} \le m \le \chi^{-1}, \\ 1/\chi^4 m^5 & \chi^{-1} \le m \le \chi^{-2}, \\ 1/\chi^{-3} m^{3/2} & m \ge \chi^{-2}. \end{cases} \tag{19}$$

When the number of bases in a strand is less than $\chi^{-2/3}$, the strands are rod like, and an m base strand in a loop has the same configurational entropy as equivalent length strands at the ends. Unfortunately experimental values of χ are not available. From the characteristic ratio of single stranded RNAs, one can estimate that $l/P \simeq 0.1$, and $\chi \simeq 0.32$ [42]. Another empirical formulation of $f(m)$ intended to bridge the region where m is small to $f(m) = 1/(m+1)^k$ for large m is given by [43]

$$f(m) = A/[(1 - C\,e^{-Bm})(m+1)^k]\,. \tag{20}$$

The originally published parameters are probably wrong. However the functional form of eq. (20) may prove useful in empirical determinations of $f(m)$. A, B and C are empirical constants.

The above estimates for the loop entropy factor assume that the single strands behave homogeneously with respect to base sequence. It is important to note that different base sequences exhibit different conformational properties as single strands [42, 44]. Poly $d(G)$, for example, forms a rigid structure which does not show any UV absorbance hyperchromicity when heated to 100°C in 0.02 M Na^+. It seems likely that $f(m)$ will have a base sequence dependence for small loops. This may be of considerable significance for the calculation of fluctuational base pair opening. It may also modify DNA melting curves. The initiation of an internal loop is governed by the product of σ and $f(1)$. A sequence dependent $f(1)$ may create an apparent sequence dependence of σ inconsistent with eqs. (10) and (11).

2.1.4. Calculating θ_{int}

Several methods have been developed to calculate θ_{int} once the model parameters and assumptions are established. A number of these procedures were summarized and compared by Gotoh [10] and Wada et al. [9]. The first method to exactly determine the partition function and concomitant thermodynamic properties of the loop entropy model was introduced in the early 1970s [13]. This method involved the generation of the statistical weights of every DNA configuration through a product of matrix operators. A different approach to the calculation of DNA melting was produced by Poland [45]. It was based on the work of Lacombe and Simha [46]. The latter approach does not calculate the partition function prior to evaluating θ_{int}. Instead, recursion formulae are derived for the conditional probabilities of the state of each base pair. The formulae arise as a consequence of the detailed balance reactions between the DNA configurations. This method has the advantage of providing the probability of being intact for each base pair without additional calculations. One may obtain a probability profile, or melting map of base pair openings along the entire DNA sequence (fig. 6). This advantage has made the Poland method the one of choice. The time required for calculating θ_{int} by either the Poland or Wartell and Montroll method is proportional to N^2 with N the number of base pairs. Fixman and Friere [40] developed an approximation to the loop entropy term $f(m)$ which reduces the calculational time to the order of N. Their approximation was based on a similar proposal by Frank-Kamenetskii and Frank-Kamenetskii [47]. $f(m)$ is substituted by a series of exponential functions

$$f(m) = \sum_{i=1}^{J} a_i \exp(-b_i m)\,, \tag{21}$$

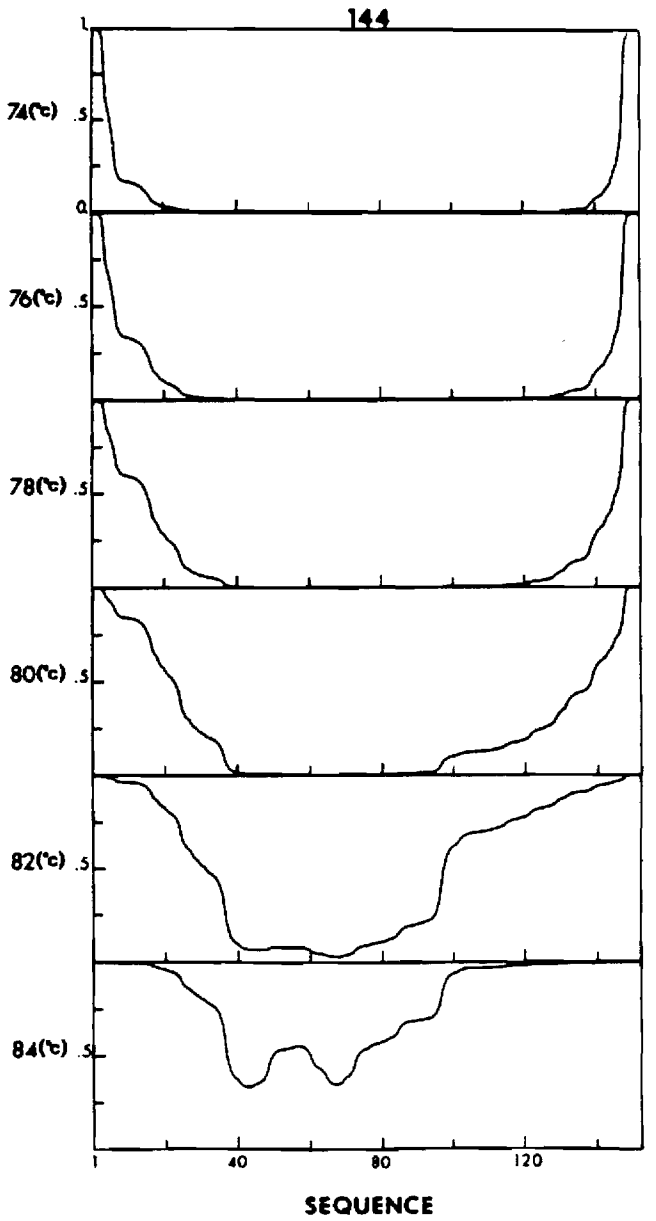


Fig. 6. Theoretical probability profiles of the 144 bp *lac* DNA at 6 temperatures. The probability of each base pair being broken ($0 \leq \theta_B \leq 1$) is plotted vs. the sequence. Dissociation of strands is not considered in this calculation.

a_i and b_i are constants determined from a numerical fit to the values of $f(m)$ for $1 \leq m \leq m_{max}$ where m_{max} is as large as seems physically relevant. For duplex DNAs with $N \simeq 5000$ the Fixman–Friere approximation greatly reduces the computer time with virtually no loss in the accuracy of θ_{int}. A J value of 9 suffices.

We will briefly review the Poland algorithm elaborating on it to include all features of the nearest neighbor model. This has not been previously presented. Two recursive formulae are required to obtain θ_{int}. One involves the conditional probability $p(l_i \mid l)$ which is the probability that the $(i+1)$ base pair is

intact ($i+1=1$) if the ith base pair is intact ($i=1$). Contrary to the conventional model description, the DNA configuration with all intact base pairs is taken as the zero of free energy. Using this calculational device, one assigns a statistical weight of 1 for an intact base pair, and $r_i = 1/s_i$ when the ith base pair is open. s_i is defined by eq. (14). Numbering the base pairs from left to right $i = 1, 2, \ldots, N$ and defining $t_i = r_i p(1_i \mid 1)$, one has the recursion relation (Poland's eq. (26))

$$t_i = r_i\left(1 + \sum_{m=1}^{N-i-1} \alpha_i(m) + \beta_i\right)^{-1}, \qquad i = N-2, N-3, \ldots, 1. \tag{22}$$

with

$$\alpha_i(m) = \sigma_{i,i+1}^{1/2} f(m) t_{i+1} \alpha_{i+1}(m-1)/\sigma_{i+1,i+2}^{1/2} f(m-1), \tag{23a}$$

$$\alpha_{N-2}(1) = \sigma_{N-2,N-1}^{1/2} f(1) \sigma_{N-1,N}^{1/2} t_{N-1}, \tag{23b}$$

$$t_{N-1} = r_{N-1}/(1 + r_{N-1}), \tag{23c}$$

and

$$\beta_i = \sigma_{i+1,i+2}^{1/2} t_{i+1} \beta_{i+1}/\sigma_{i+1}^{1/2}, \tag{24a}$$

$$\beta_{N-2} = r_N t_{N-1}/\sigma_{N-2,N-1}^{1/2}. \tag{24b}$$

The conditional probabilities are obtained through a recursive sweep using eq. (22).

The second recursive formula involves the unconditional probability, $p(1_i)$, that the ith bp is intact. Poland's recursive relation modified to include sequence dependent σ's is given by

$$p(1_{i+1}) = p(1_i)\gamma_i/\sigma_{i,i+1}^{1/2} + \sum_{j=1}^{i-1} p(1_j)\mu_i(j) + p(1_i)t_i/r_i, \tag{25}$$

where

$$\gamma_i = t_i \gamma_{i-1}, \tag{26a}$$

$$\gamma_1 = t_1, \tag{26b}$$

and

$$\mu_i(j) = \sigma_{i,i+1}^{1/2} f(i-j) t_i \mu_{i-1}(j)/\sigma_{i-1,i}^{1/2} f(i-j-1), \qquad j = 1, 2, \ldots, i-2, \tag{27a}$$

$$\mu_i(1) = t_1/r_1, \tag{27b}$$

$$p(1_1) = \left[1 + \sum_{n=1}^{N-1} \prod_{j=1}^{n} t_j\right]^{-1}. \tag{28}$$

Strand separation is excluded in this formulation and is considered later. $p(1_i)$ is calculated for all $i = 1, \ldots, N-1$, by using eq. (25). The average fraction of intact base pairs for DNAs with at least one

base pair is obtained from

$$\theta_{\text{int}} = \sum_{i=1}^{N} p(l_i)/N. \tag{29}$$

Calculation of θ_{int} by the above algorithm provides an exact result. An approximate method which provides insights into the melting process was presented by Azbel [41]. Unlike prior approximate methods [4, 11] which averaged together blocks of base pairs to simplify calculation of the entire partition function, Azbel simplified the partition function by considering only the configuration of lowest free energy at each temperature. A procedure was developed which determined the lowest free energy state given the DNA sequence, the temperature, and model parameters. Since experimental estimates on σ implied a long correlation length in the transition region, this method considered DNA melting to occur sequentially in steps. The duplex configuration is the lowest energy state below the transition. It dominates the partition function until a certain temperature, at which point a region of 100 bp or so opens up. This new configuration replaces the former state as the lowest free energy state, which in turn is the dominant state until a higher temperature. Only five types of local melting transitions can occur. These are concisely described by Gotoh [10] in his review of Azbel's work, and a similar description is given in fig. 7. The transition temperature, T_m, of a specific region depends on the region's internal stability and the net change in junctions induced by the region melting. If we include nearest neighbor interactions, the Azbel method yields for a region of length L

$$T_m - \left[X_{AT}T_{AT} + (1 - X_{AT})T_{GC} + \sum_{\{MN\}} X_{MN}\,\delta G_{MN}\right] \Big/ \left(1 + \frac{R}{L\,\Delta S}\ln\frac{\sigma_b}{\sigma_a}\right), \tag{30}$$

X_{AT} is the fraction of A·T pairs in the region. The sum {MN} is over the ten different base pair doublets. X_{MN} is the number of MN doublets divided by L. σ_b divided by σ_a represent the ratio of $f(m)$'s and σ's which occur before and after the region opens. For a type I transition,

$$\sigma_b/\sigma_a = \sigma_{i,i+1}^{1/2}/\sigma_{i+L,i+L+1}^{1/2}. \tag{31a}$$

For a type II transition,

$$\sigma_b/\sigma_a = 1/f(L)\sigma_{i,i+1}^{1/2}\sigma_{i+L,i+L+1}^{1/2}. \tag{31b}$$

Normally when a single cooperativity parameter is used one does not subscribe the σ parameters, and $\sigma^{1/2}$ is unnecessary.

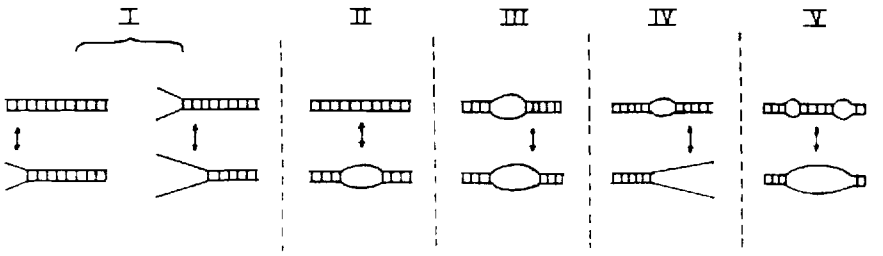


Fig. 7. The five types of local melting transitions which may occur in a linear DNA.

Equations (30) and (31) illustrate that if nearest neighbor interactions are included completely, they affect both the internal stability of the region (the numerator of eq. (30)) and the free energy term for duplex-single strands junctions (the denominator of eq. (30)). Gotoh compared the Azbel method to an exact calculation, and found significant differences. The omission of the many microstates of the partition function are apparently important. Determination of the exact locations and sizes of melting regions can be ambiguous [48]. This does not necessarily invalidate the principle of the method, but suggests that higher energy configurations must be included. The methodology required to include such terms has been discussed [41].

2.2. *DNA strand dissociation; calculating* θ_{ext}

Several authors have described models which include effects of duplex to single strands dissociation [6, 11]. The approach described here is a modification of the derivation by Poland and Scheraga [11]. It considers the influence of the internal states of the DNA on its external degrees of freedom. For DNAs less than about 300 bp in length, dissociation has a significant influence on the melting curve. Figure 8 demonstrates this for a 95 bp DNA. The initial 20% of the transition is governed by θ_{int}. Beyond this stage, strand dissociation rapidly drives the melting curve to completion. If one could covalently join the two strands to prevent dissociation, the resulting curve should behave as θ_{int} throughout. The dissociation of duplex DNA to single strands is described as

$$C_2 \rightleftarrows 2C_1 , \tag{32}$$

where C_2 represents the two stranded DNA molecule and C_1 represents a single strand. We assume the two complementary strands behave as equivalent polymers. The equilibrium dissociation constant for

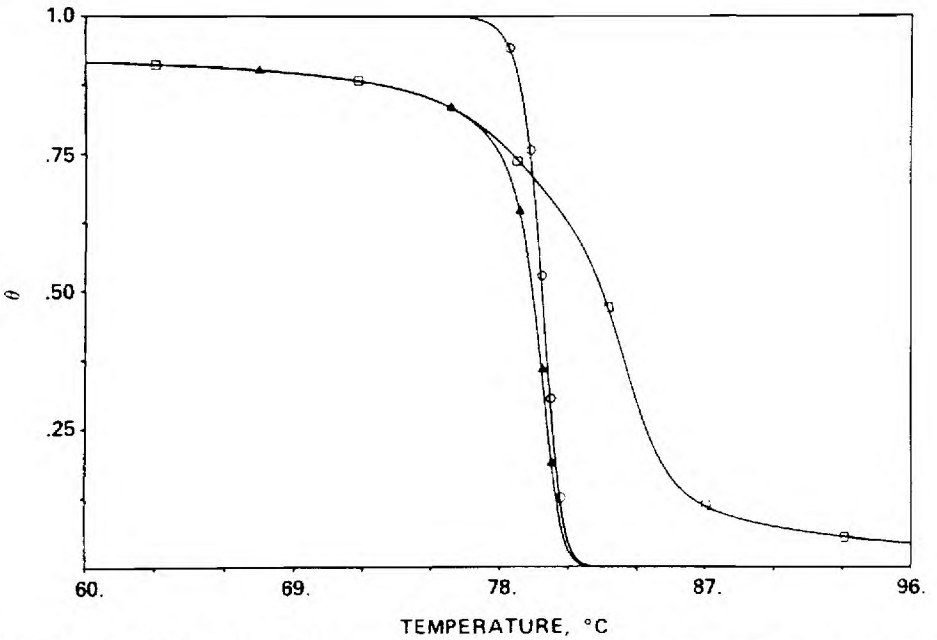


Fig. 8. Theoretical melting curves of 95 bp *lac* DNA. The overall melting curve, $\theta = \theta_{int}\theta_{ext}$ is given by (▲), θ_{ext} is given by (○) and θ_{int} is given by (□).

eq. (32) is given in terms of the concentrations of C_1 and C_2

$$K_D = C_1^2/C_2 . \tag{33}$$

Employing eq. (33) and conservation of the total DNA concentration, $C_T = C_1 + 2C_2$, we can express θ_{ext} as

$$\theta_{ext} = 2C_2/C_T = \frac{4C_T/K_D + 1 - [1 + 8C_T/K_D]^{1/2}}{4C_T/K_D} . \tag{34}$$

In melting experiments, the DNA concentration is low, (~60 μM bp) therefore the DNA is assumed to behave as an ideal solute. For ideal solutions, the solute molecules comprising the ensemble may be treated as if they do not interact with one another except for collisions [49, 50]. Therefore, K_D can be expressed as a quotient of molecular partition functions for the duplex and single stranded species [11]. We assume these partition functions can be written in terms of external and internal degrees of freedom of the two DNA species, i.e., the product of external and internal partition functions.

$$K_D = \frac{[Z(C_1)]^2}{Z(C_2)} = \frac{[Z_{ext}(C_1)]^2[Z_{int}(C_1)]^2}{Z_{ext}(C_2)Z_{int}(C_2)} . \tag{35}$$

Letting

$$\beta = Z_{ext}(C_2)/Z_{ext}^2(C_1) , \tag{36}$$

leaves

$$K_D = Z_{int}^2(C_1)/\beta Z_{int}(C_2) , \tag{37}$$

where $Z_{int}(C_1)$ and $Z_{int}(C_2)$ include the vibrational and electronic energies and $Z_{ext}(C_1)$, $Z_{ext}(C_2)$ consider the translational and rotational energies. $Z_{int}(C_1)/Z_{int}(C_2)$ is obtained in the process of evaluating θ_{int}.

Each of the external partition functions in eq. (36) can be expressed as a product of the translational and rotational partition functions of the molecule in question. The mass dependence, and hence N dependence of translational and rotational partition functions of polyatomic molecules [51] suggests that Z_{ext} depends on the DNA base pair length N, as

$$Z_{ext}(C_1) \propto N^{\alpha_1} , \tag{38a}$$

$$Z_{ext}(C_2) \propto N^{\alpha_2} , \tag{38b}$$

and therefore

$$\beta = KN^{\alpha_2 - 2\alpha_1} = KN^{\alpha} , \tag{39}$$

where K includes all factors independent of N. α_1 and α_2 combine the N dependence of translational

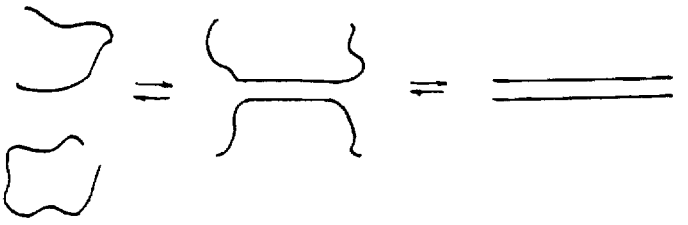


Fig. 9. Sketch illustrating the overall shape of a DNA in several stages of the melting transition.

and rotational degrees of freedom. Translational partition functions vary as $N^{3/2}$. Rotational partition functions depend on the product of the principal moments of inertia of the molecules. For a uniformly dense globular or 'sphere-like' molecule Z_{rot} goes like $N^{5/2}$. A rod-like molecule has Z_{rot} proportional to $N^{7/2}$.

The choice of α defining the N dependence of β in eq. (39) depends on the global conformation assumed by duplex helix and coiled single strands in different regions of the transition. Thus, if dissociation is modeled as an equilibrium of one rod-like molecule dissociating to two sphere-like molecules $\alpha = -3$. Modeling the equilibrium as one sphere-like molecule dissociating to two sphere-like molecules, $\alpha = -4$. DNA molecules 100–500 bp in length can not be realistically represented by either of the models given above. Comparison of theory and experiment have verified this and suggested a simple first-order improvement [23]. Dissociation was represented by a combination of the two models. In the early stages of a transition a rod unwinding to two sphere-like molecules was considered. In later stages of the transition the average duplex DNA has a large portion of single stranded regions and a sphere-like molecule dissociating to two sphere-like molecules was assumed (fig. 9). This configurationally variable process was modeled by making α depend on θ_{int},

$$\alpha = a + b(1 - \theta_{int}), \tag{40}$$

where a and b are constants which allow α to vary from -3 to -4 through the DNA transition region.

Although the formulation of β from eqs. (39) and (40) have provided a successful description for the melting curves of short DNAs, it is acknowledged that rods and spheres are oversimplified models for DNA. Hydrodynamic and light scattering measurements indicate that DNAs are better represented as a worm-like chain [14, 52]. These studies also demonstrate that the flexibility of DNA is influenced by the ionic strength of the solvent. A reduced ionic strength increases phosphate repulsion along the backbone. This can modify single stranded conformations. Another factor which can effect β is the DNA sequence. Base sequences will determine the extent of stacking rigidity of a strand and intrastrand structures. Results described in section 3 demonstrate that both base sequence and ionic strength influence β. The use of eq. (40) can be justified on empirical grounds for heterogeneous sequence DNAs [53].

3. Comparison of theory with experiment

3.1. Melting curves of DNAs longer than 1000 bp

A number of comparisons have been made between theoretically calculated differential melting curves and their experimental counterpart. The first comparisons were made with DNAs from ϕX174

phage DNA [20, 54, 25]. Three molecules were examined, the whole ϕX174 DNA (5386 bp) and two *Hpa* II restriction endonuclease fragments of this DNA designated Y1 (2748 bp) and Y2 (1695 bp). Initial theoretical predictions employed two stability parameters s_{AT} and s_{GC} and one σ. Agreement between theory and experiment was not observed. This lack of agreement may, however, be due to causes other than the inadequacy of the theory.

The studies on the whole ϕX174 DNA were done with samples assumed to be randomly broken once per molecule. Closed circular ϕX174 DNA was treated with enzymes thought to make random nicks [20, 25]. Recent studies have shown that this cleavage is not random [55, 56]. The linear Y1 and Y2 fragments do not suffer from this problem. However a different concern of a more general nature must be considered for these fragments. This is the question of whether the experimental curves are at equilibrium. Is it valid to compare them with theoretical equilibrium predictions?

Experimental melting curves are generally obtained by heating DNA solutions at a rate of 0.1–0.2°C/min. This progressive heating method – rather than temperature step equilibrium – reduces the time of an experiment to about 6–7 h. This minimizes the probability of generating single stranded breaks in the DNA due to extended exposure of the sample to high temperatures. In order for a melting curve to be an equilibrium one, the relaxation times for DNA unwinding must be less than the time allowed by the heating rate. Several workers have reported that hysteresis occurs in DNA melting curves in low ionic strength buffers [57–59]. Partially denatured DNA when cooled at the same rate as the heating rate show irreversibility in their melting curves. This irreversibility is clear evidence for non-equilibrium melting. The influence of this non-equilibrium on the forward melting curve is discussed later. Perelroyzen et al. [59] has shown it occurs in the popular buffer containing 0.015 M sodium chloride and 0.0015 M sodium citrate (0.1 SSC, 0.0195 M Na^+). These authors also demonstrated that in a buffer containing 0.2 M Na^+, DNA melting curves were reversible prior to final strand separation. Reversibility for DNA melting curves has been observed by us in 0.1 M Na^+ (see section 3.4). Blake and Haydock have also observed reversible behavior for a solvent containing 0.08 M Na^+ [60].

The above results suggest that the least controversial comparisons of theory with experiment be made with solvents containing at least 0.08 M Na^+. Unfortunately the initial theory–experiment comparisons with the Y1 and Y2 DNAs were made in 0.0195 M Na^+.

A survey of the literature finds two instances where melting curve data on the Y1 and Y2 fragments were obtained in high salt. Figure 10 shows melting curves of the Y1 and Y2 DNAs in the 1.0 SSC buffer (nominally 0.2 M Na^+). Tachibana et al. [61] determined the curves for both fragments, and Perelroyzen et al. [59] obtained data for the Y2 DNA. The peaks of the last transition of the two data on the Y2 DNA differ by 0.4°C. The difference in melting shapes is considerably larger. Part of the difference may be due to different normalization procedures used to relate absorbance to θ. The integrated intensities of the two curves should be the same, yet they are not. Tachibana et al. calculated the derivative curve by fitting a cubic polynomial to each segment containing 15 consecutive data points, and differentiating. Perelroyzen et al. obtained their differential melting curve by digital differentiation followed by a smoothing with a Gaussian function. The theoretical curves shown in fig. 10b employed two stability parameters s_{AT} and s_{GC} and one σ. Values for the parameters are given in the fig. 10b legend. The theory curve for the Y2 DNA is mildly similar to experiment. The one for the Y1 DNA is clearly outside of experimental error. Other workers have also noted a lack of agreement between theory and experiment [25, 61]. We have not allowed ourselves the liberty of varying the parameters for each DNA fragment separately. One can improve the fit for the Y1 DNA, but no single set of parameters (using two s's and one σ) was found which fit both melting curves.

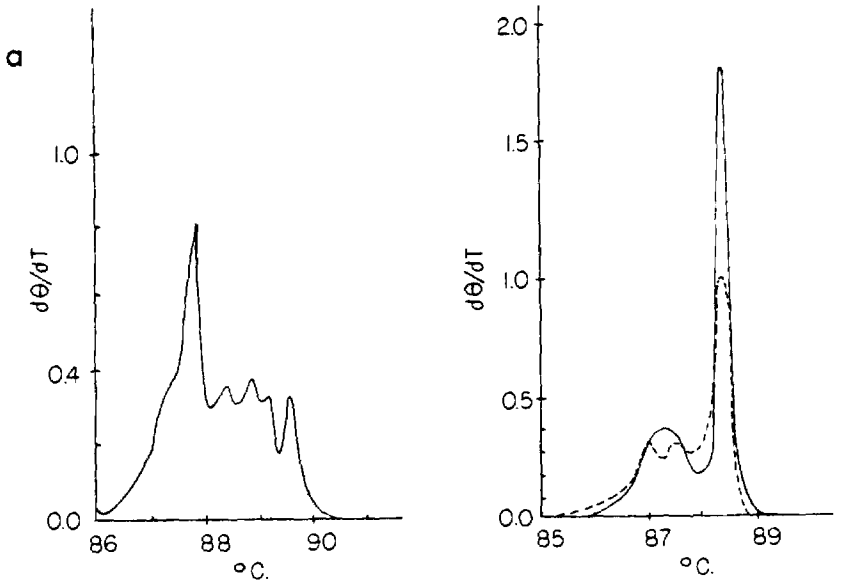


Fig. 10a. Experimental differential melting curves in 195 mM Na^+ of the Y1 DNA fragment (on the left), and the Y2 DNA (on the right). The two curves for the Y2 DNA correspond to data obtained by Tachibana et al. [61] (———), and Perelroyzen et al. [59] (– – – –).

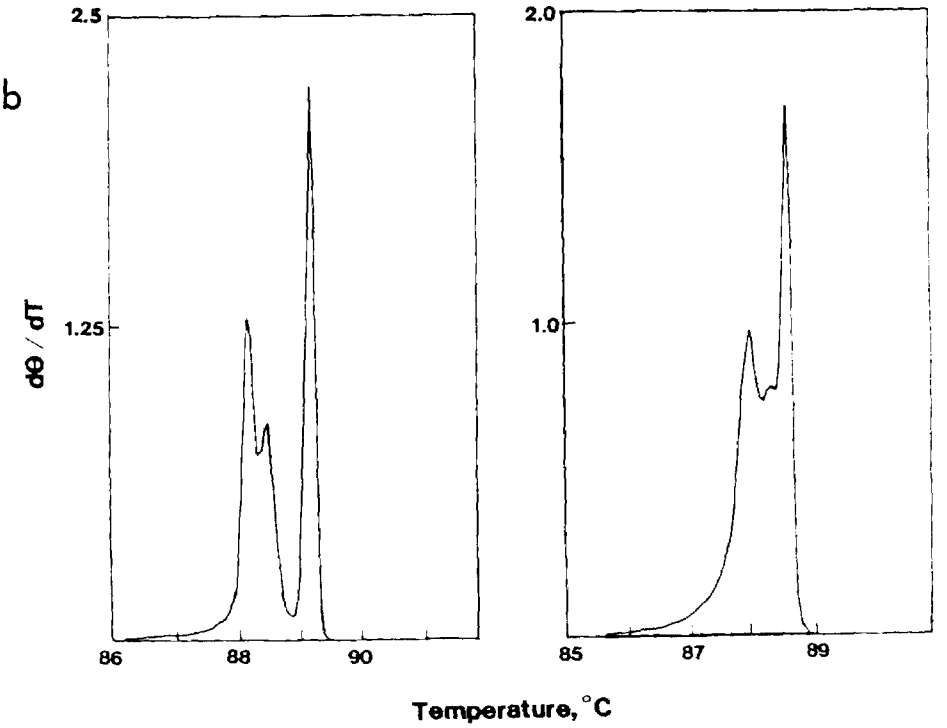


Fig. 10b. Theoretical melting curves of the Y1 and Y2 DNAs. Fixman–Friere algorithm was employed with the parameters: $T_{AT} = 70.5°C$, $T_{GC} = 111.2°C$, $\sigma = 10^{-6}$ and $f(m) = 1/(m+1)^k$, $k = 1.7$.

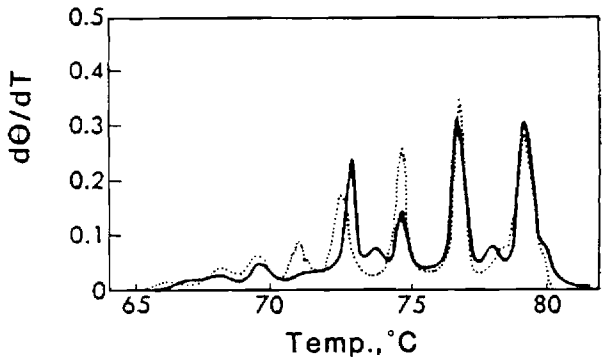


Fig. 11. Theory–experiment comparison for pBR322 DNA in 0.0195 M Na^+ obtained by Gotoh (10). Dotted line is experiment.

Gotoh and Tagashira [27] incorporated the effect of nearest neighbor interactions in order to improve theory–experiment agreement (section 2.1.2). Ten s values were determined by a least squares approach similar to that described earlier (table 1). The Fixman–Friere procedure was employed to predict differential melting curves of 10 DNA molecules which included the Y1 and Y2 fragments of ϕX174 DNA and eight restriction endonuclease fragments from the fd virial DNA. The agreement was considerably better than results obtained with the model using two s values. This in itself is not surprising since the experimental data was used to empirically determine the ten additional parameters. However independent comparisons with other data suggest significance to their result. The evaluated parameters predict with good accuracy [10] many of the peaks in the melting curves for pBR322 DNA (4362 bp) and a T7 phage (1459 bp). Figure 11 reproduces the comparison for the pBR322 DNA. The only qualm with the above findings is that all of the experimental curves considered were obtained in 0.1 SSC or about 0.02 M Na^+. This solvent produces non-equilibrium melting behavior. Perelroyzen et al. [59] and Gotoh [10] qualitatively examined how non-equilibrium conditions might cause experimental profiles to deviate from equilibrium profiles. Although it was concluded that deviations from the actual equilibrium melting curve will occur, a quantitative estimate of this deviation was not made. If the experimental melting curves deviate significantly from equilibrium, then the agreement between theory and experiment observed by Gotoh is difficult to interpret. The evaluated cooperativity parameters and stability parameters may reflect an empirical fit, rather than thermodynamic parameters transferable to other DNAs and other situations (e.g., low temperatures). It should be noted that the fit of theory with experiment for the whole fd virial DNA (6408 bp) was not as good as with fd DNA fragments [10]. Anshelevich et al. [62] have derived equations to estimate how much a given kinetic heating rate will shift the observed T_m of a cooperatively melting region from its equilibrium value. We will review this analysis.

3.2. The effect of non-equilibrium conditions on the melting curves of long DNAs

Let us assume that when a DNA region unwinds under equilibrium conditions it unzips within a temperature interval ΔT about its T_m. If the heating rate in an experiment is v (°C/s), then in order for equilibrium conditions to hold, the relaxation time for unwinding at $T = T_m$ must obey

$$t_{rel}(T_m) < t_{exp} = \Delta T/v\,. \tag{39}$$

Observations of equilibrium melting curves show that the width of peaks corresponding to 100–300 bp regions is $\Delta T \simeq 0.3$°C. Assuming $v = 0.002$°C/s, the relaxation time $t_{rel}(T_m)$ must be less than 150s. If $t_{rel}(T_m)$ is greater than this value, the DNA region will melt at a temperature T_k which is higher than T_m. T_k is defined by $t_{rel}(T_k) = t_{exp}$. At T_k the temperature driven unwinding rate for the region catches up with the experimental heating rate. Anshelevich et al. [62] estimated $t_{rel}(T)$ by examining the overall kinetics of unwinding from both ends of a homopolymer sequence of N base pairs. Their result is

$$t_{rel}(T) = \frac{1}{2k_f} \frac{s^{N+3}}{(s-1)^2[N(s-1)-2]}, \tag{42}$$

k_f is the elementary rate constant for forming an intact base pair onto a helical region. It has been estimated at $10^6\,s^{-1}$ [63]. s is the familiar stability parameter. The temperature dependence of the relaxation time is dominated by the s^N term. The relaxation time is a very sensitive function of temperature. Equation (42) is a reasonable approximation of unwinding of a heterogeneous sequence of A · T and G · C pairs only if the base pair distribution is homogeneous. Actual relaxation times may be longer.

Experimental and theoretical studies by Wada et al., Perelroyzen et al. and Michel [9, 59, 64] showed that irreversibility in 0.1 SSC occurs after the initial part of the melting curve. These studies indicated that the onset of non-equilibrium behavior occurs due to the unwinding of DNA region in the type IV and V transitions of fig. 7. This was initially proposed by Hoff and Roos based on the work of Spatz and Crothers [65, 66]. An estimated of $t_{rel}(T_m)$ and the difference between T_k and T_m for a type V transition is determined by evaluating $s(T_k)$ and $s(T_m)$. $s(T_k)$ is obtained by equating $t_{rel}(T_k)$ to t_{exp}. $s(T_m)$ is obtained by equating the equilibrium constant for this transition, K_V, to 1, or

$$K_V = q\sigma f(N_a) f(N_c)[s(T_m)]^{N_b} / f(N_a + N_b + N_c) = 1\,, \tag{43}$$

where σ and $f(N)$ are the cooperativity factor and loop entropy factor. The term q is a degeneracy factor. It compensates for the number of configurations of roughly equivalent free energy which result from an exact calculation of a type V transition. Based on exact calculations Anshelevich et al. find $q \sim 10^2$. Assuming $f(N) \sim N^{-k}$ with $k = 1.5$ and $N_a \simeq N_b \simeq N_c \simeq N$, $s(T_m)$ was determined for a σ value estimated for an ionic strength of 0.02 M Na^+ ($\sigma \sim 10^{-7}$). Our results described later indicate σ is larger. The relaxation time for unwinding $N = 300$ bp with these assumptions is about $t_{rel}(T_m) \simeq 10^3$ s. This is consistent with the non-equilibrium behavior observed for this solvent. From $s(T_k)$ and $s(T_m)$, the authors obtained

$$\delta T_{mk} \equiv T_k - T_m = \frac{RT_\infty^2}{N\,\Delta H}\left[\ln\left(\frac{N^{k+2}}{\sigma q 3^k}\right) - \kappa\right], \tag{44a}$$

with

$$\kappa = 1.15 \ln(2k_f \Delta T / v) + 5.5\,. \tag{44b}$$

T_∞ is the melting temperature of an infinitely long DNA determined by $s = 1.0$, ΔH is the enthalpy change corresponding to base pair formation. For a solvent of 0.02 M Na^+, one finds for $N = 300$, $\delta T_{mk} \simeq 0.2$°C. Not considered in this analysis is the effect of the non-equilibrium behavior of one region on the melting of adjacent regions [67].

The above findings indicate that a melting curve taken in 0.02 M Na^+ will differ significantly from equilibrium when type IV and V transitions occur. If these transition types do not occur then a DNA melting curve will be close to equilibrium. This situation together with the large number of potential configurations in a long DNA muddles theory–experiment comparisons. Real discrepancies between theory and experiment are difficult to distinguish from deviations caused by the non-equilibrium melting of DNA regions. The calculated non-equilibrium shift of $\delta T_{mk} = 0.2°C$ for 0.02 M Na^+ is likely to be a lower limit. Gotoh and Tagashira's work clearly demonstrated that stacking heterogeneity can be important. The difficulties caused by non-equilibrium melting and the complexity of long DNA melting curves make the generality of their parameters uncertain. As described in the following section, they do not predict the melting curves of short DNAs taken in 0.1 M Na^+.

3.3. Melting curves of DNAs shorter than 600 bp

The development of techniques to isolate short DNA restriction fragments [68, 69] made it possible to carry out melting studies on homogeneous molecules 100–600 bp long. A study of DNAs in this size range has several advantages over kilo base pair length molecules. Experimental melting curves have only one to four subtransitions. This observation coupled with the short DNA lengths simplifies the theory–experiment comparisons. The high free energy cost in forming internal loops means that end melting will govern the unwinding process for a quasi-random sequence DNA. Theoretically predicted melting curves will be largely insensitive to the parameters σ and $f(m)$. One can concentrate attention on the effect of stacking heterogeneity on s values. On the other hand short DNAs with thermally stable regions at the ends and a thermally unstable region in the middle can be used to evaluate $\sigma f(m)$ for small m. Theoretical calculations have shown that the melting curve of a 157 bp DNA with this type of sequence is sensitive to the σ and $f(m)$ values [53]. One disadvantage of the short DNAs is that strand dissociation cannot be neglected (section 2.2). This adds three additional parameters to the theoretical calculation. Results described below indicate that these parameters are generally not sensitive to DNA sequence. The three added parameters (eqs. (39) and (40)) do not diminish the sensitivity of the melting curves to the s values. From a more positive outlook the short DNAs provides a way of evaluating dissociation parameters in this size range.

The first comparison of DNA melting theory with experimental transitions of short DNA restriction fragments was made with fragments 95, 144, 203 and 301 bp long [23]. Figure 12 displays the comparison. The experimental transitions of the 95, 203 and 301 bp DNAs were obtained by Hardies et al. [70]. The denaturation curve of the 144 bp DNA was obtained in this laboratory. The solvent employed contained 0.1 M NaCl + 1 mM potassium phosphate and 0.1 mM EDTA. Although the reproducibility of the T_m's for a given DNA was ±0.5°C, the reproducibility of the shape and internal temperature differences within a profile were much better. Two differential melting curves for the 301 bp DNA are shown in fig. 12. The theoretical results shown in fig. 12 employed the model with two stability parameters s_{AT} and s_{GC} and one σ. As expected, variations in σ had little influence on the predicted curves ($2 \times 10^{-6} \leq \sigma \leq 1 \times 10^{-4}$). It should be mentioned that the theoretical curves were translated by 0.3–0.9°C to match the location of the major theory peak with experiment. These shifts were within experimental uncertainty of the T_m's. A second study by Hillen et al. [72] extended the theory–experiment comparison to four more fragments 80, 101, 188 and 219 bp in length. The agreement between experimental and theoretical curve shapes and T_m values was similar to that observed in fig. 12. Figure 13 displays the melting curves of these four additional DNAs. The G·C content of the eight DNAs varied from 44.4 to 64.4%.

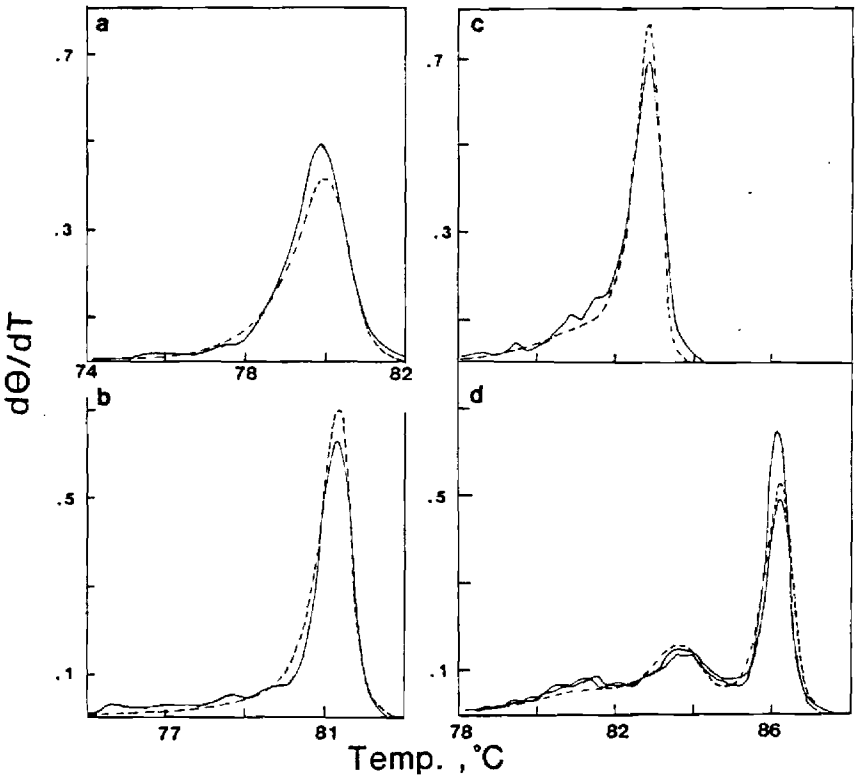


Fig. 12. Comparison of theoretical (-----) and experimental (———) DMCs for four *lac* operon fragments as reported in ref. [23]; (a) 95 bp, (b) 144 bp, (c) 203 bp, (d) 301 bp. Two experimental curves are shown for the 301 bp DNA. Experimental curves for the 95, 203 and 301 pb DNAs were obtained by Hardies et al. [70]. Theoretical parameters were $T_{AT} = 64.3°C$, $T_{GC} = 107.2°C$, $\sigma = 4.5 \times 10^{-5}$, $\Delta H_{AT} = -8.5$ kcal/mole, $\Delta H_{GC} = -9.4$ kcal/mole, $f(1) = 1$, $f(2) = 0.95$, $f(3) = 0.90$, $f(m) = 1/(1 - 1.38\exp(-0.1m)(m+1)^{1.7}$ for $m \geq 4$. Dissociation parameters (eqs. (39) and (40)) were $K = 5000$, $a = -2.8$, $b = -3.2$.

The agreement observed for the eight DNA fragments in 0.1 M Na^+ indicated that the model deduced for strand dissociation was generally correct. The shapes of the final peaks of the DMC were highly sensitive to the dissociation parameter β [53]. Other aspects of the comparisons suggested that further studies with additional DNA fragments would be illuminating. The good predictions obtained with only two stability parameters, s_{AT} and s_{GC}, indicated that the heterogeneity of the stacking free energies (δG_{MN}) is not large. We have examined the four different sets of δG_{MN}'s in table 1 with the 144, 203 and 301 bp DNAs. Only set 1 gave predictions which are in fair agreement with experiment. The inclusion of this set of ten s values shifted the predicted T_m's by about ±0.5°C. Alterations in the peak heights and internal temperature differences of subtransitions were not large. Parameter sets 2, 3 and 4 of table 1 distorted the predicted curves far beyond experimental error. Figure 14 shows an example of this for the 301 bp DNA using the Gotoh and Tagashira parameters. The shape of the DMC is not produced and the T_m is predicted to be 3°C too low. Similar poor agreement was observed for the other DNAs. Even if adjustments were made in β, the Gotoh and Tagashira parameters do not agree with the short DNA melting curves.

Additional experimental melting curves were obtained from a series of five *Hae* III restriction fragments and two *Hinf* I fragments from the plasmid pBR322. These DNAs varied in length from 154 to 587 bp. Table 2 lists the lengths, % G · C content and the restriction endonucleases generating these

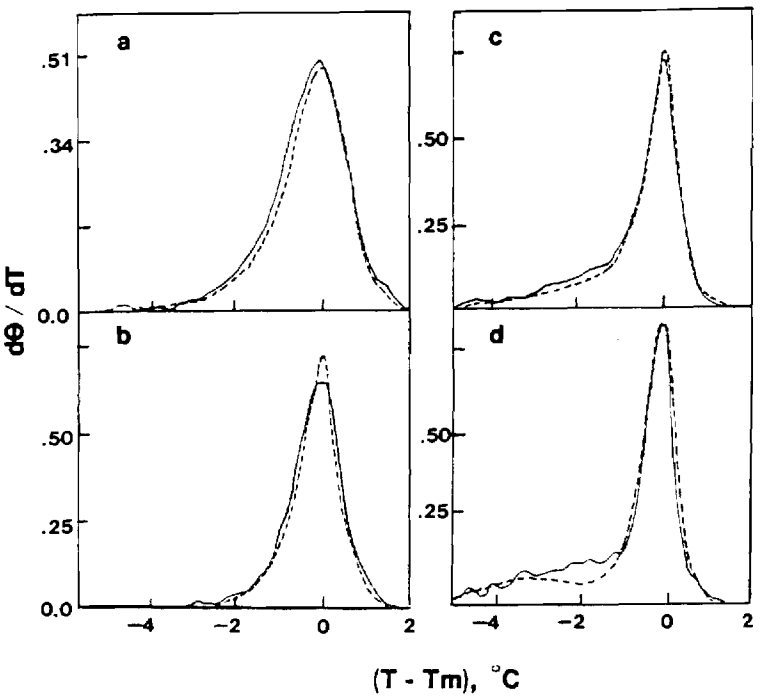


Fig. 13. Theoretical (-----) and experimental (———) DMCs of four *lac* DNAs as reported in ref. [71]; (a) 80 bp, (b) 101 bp, (c) 188 bp, and (d) 219 bp. Theory parameters are given in fig. 12.

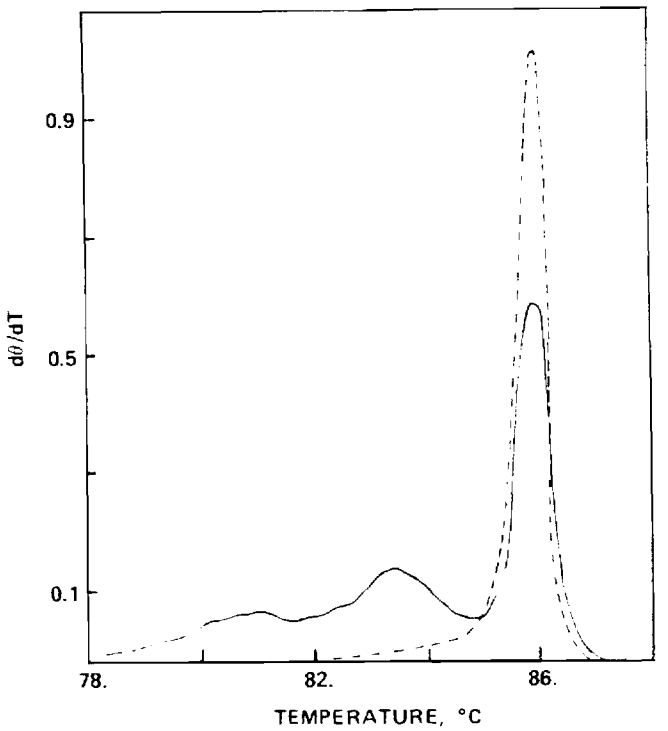


Fig. 14. Experimental DMC of 301 bp *lac* DNA (———) and theoretical curve predicted with Gotoh and Tagashira δG_{MN} parameters, set 3 in table 1 (-----). Other theory parameters as in fig. 12. The two curves are aligned for comparison. T_m^{th} was 3°C less than T_m^{ex}.

Table 2
pBR322 DNA fragments isolated for melting curve studies

Fragment length (bp)	% GC	Enzyme sites of ends
587	49.0	Hae III
516	44.0	Hinf I
434	57.8	Hae III
267	50.9	Hae III
234	59.0	Hae III
192	47.4	Hae III
154	62.4	Hinf I

fragments. The isolation and purification of these fragments will be described elsewhere (Benight, Abhiraman and Wartell, manuscript in preparation). The appendix describes the instrument and data analysis procedures used to obtain the differential melting curves.

Predicted melting curves for the pBR322 DNAs were initially made using the parameters employed for the eight *lac* DNAs. These parameters are described in fig. 12. Agreement between theory and experiment was good, but the errors were larger than before. T_m differences of 0.7–1.3°C were noted in several cases, and the melting profiles were not always within expected errors. Rather than try a least squares approach to all of the data, small adjustments were made in the dissociation parameters, and a few non-zero δG_{MN} values were employed to fit the observed curves for the 144 bp, 192 bp and 154 bp DNAs. These admittedly arbitrary adjustments resulted in a parameter set which produced surprisingly good fits to nearly all of the pBR322 DNAs. Table 3 lists the parameters employed and the predicted and observed characteristics for the seven pBR322 and the 144 bp *lac* DNA melting curves. These eight DNAs were examined and analyzed using the same procedures (see appendix). Figures 15, 16, 17 and 18 compare the theoretical melting curves with experiment for the 144 bp *lac* DNA and the seven pBR322 DNAs. With the exception of the 516 bp DNA (fig. 18) the agreement between theory and

Table 3
Experimental and theoretical melting curve characteristics of DNA fragments (solvent conditions: 0.105 M Na^+). Theoretical parameters: T_{AT}, T_{GC}, ΔH_{AT}, ΔH_{GC}, $\tilde{\sigma}$, $f(m)$ are given in fig. 12 caption. δG_{MN} values: $\delta G_{AA} = -150$, $\delta G_{GC} = -100$, $\delta G_{CG} = 65.0$, all other $\delta G_{MN} = 0$. Dissociation parameters: $K = 5000$, $a = -2.7$, $b = -3.0$

Fragment (No. bp)	T_m(°C)		$(d\theta/dT)_{max}$	
	Expt.	Theory	Expt.	Theory
lac 144	81.3	81.9	0.82	0.66
pBR322 154	91.1	91.2	0.88	0.86
pBR322 192	80.7, 84.3	80.4, 84.0	0.17, 0.29	0.17, 0.41
pBR322 234	89.1	88.9	1.25	1.42
pBR322 267	85.0	85.6	1.30	1.35
pBR322 434	88.9	89.0	0.92	0.97
pBR322 516	78.7, 81.6, 85.0	79.8, 85.0	0.17, 0.1, 6.2	0.38, 0.92
pBR322 587	80.5, 84.2	80.8, 84.4	0.18, 0.60	0.17, 0.81

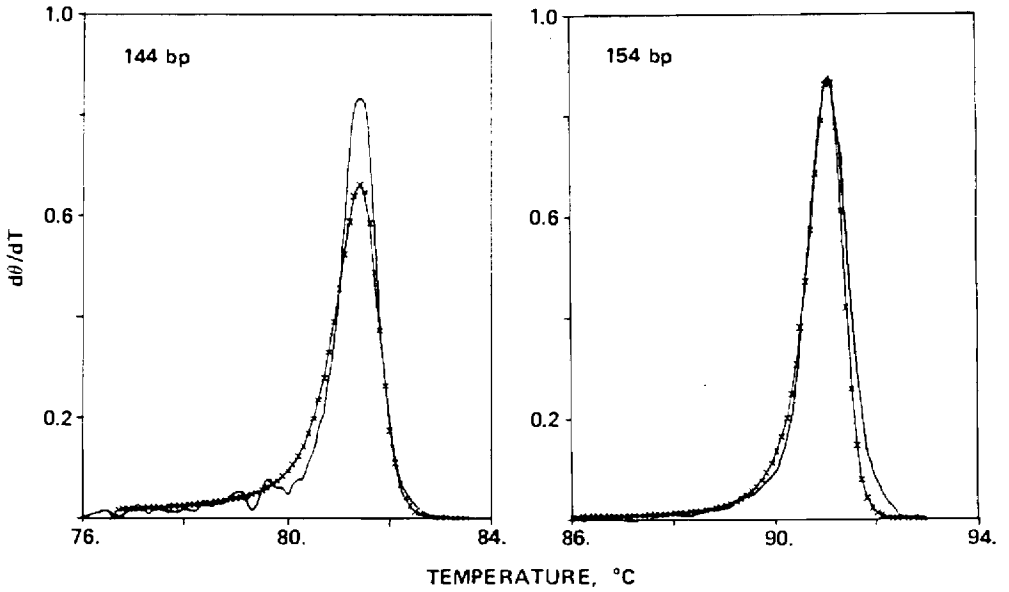


Fig. 15. Theoretical (×) and experimental (———) DMCs of 144 bp *lac* DNA and 154 bp pBR322 DNA. Characteristics of curves and theoretical parameters given in table 3. The peak height of the 144 bp experimental DMC differs from fig. 12 due to a different analysis procedure (see appendix).

experiment is very close to experimental uncertainty. The predicted melting curve of the 516 bp DNA was extensively analyzed in an attempt to find a set of parameters which gave better agreement with experiment. These attempts were unsuccessful. An examination of the base pair sequence of this DNA revealed a possible explanation for this failure. Three inverted repeat sequences occur in the 516 bp DNA. These sequences, located 220, 278 and 376 bp from one end, can form 'hairpin' or stem-loop structures with at least 9 to 11 bp in the stem regions, and 4 to 7 bases in the loops (fig. 19). Moreover

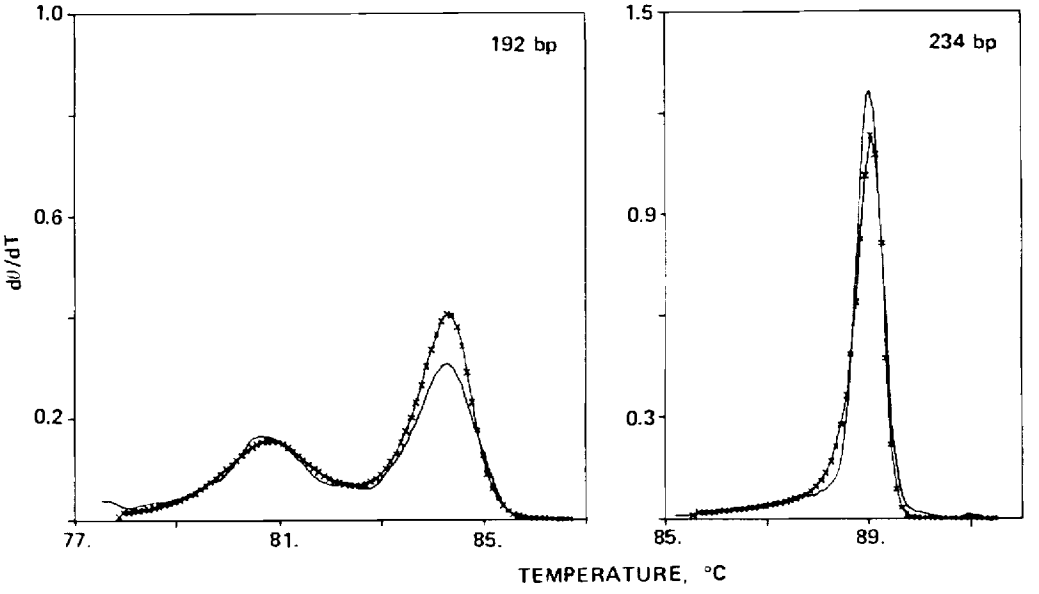


Fig. 16. Theoretical (×) and experimental (———) DMCs of 192 bp and 234 bp DNAs. See table 3 for further information.

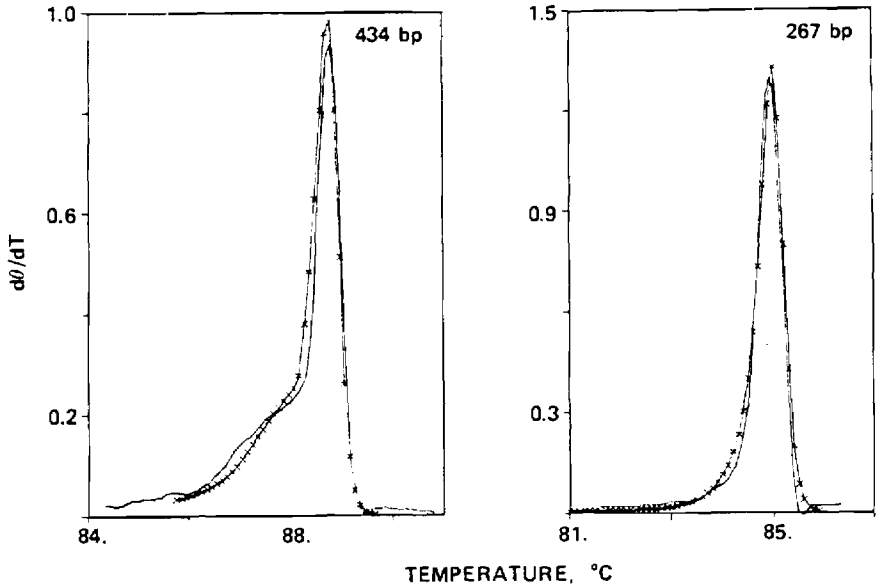


Fig. 17. Theoretical (×) and experimental (——) DMCs of 267 bp and 434 bp DNAs. See table 3 for further information.

experimental studies by Lilly [56] have shown that these three sites are the only ones in the entire pBR322 DNA which are induced into stem-loop structures by negative supercoiling. It seems plausible that these stem-loop structures occur during the melting of the 516 bp DNA fragment. If this is true then it is not surprising that the theory is unable to predict the 516 bp DNA DMC. The model does not consider the possibility of stem-loop structures. Future theoretical investigations should be capable of verifying this explanation.

The new parameters employed in table 3 were employed to recalculate the melting curves of the *lac*

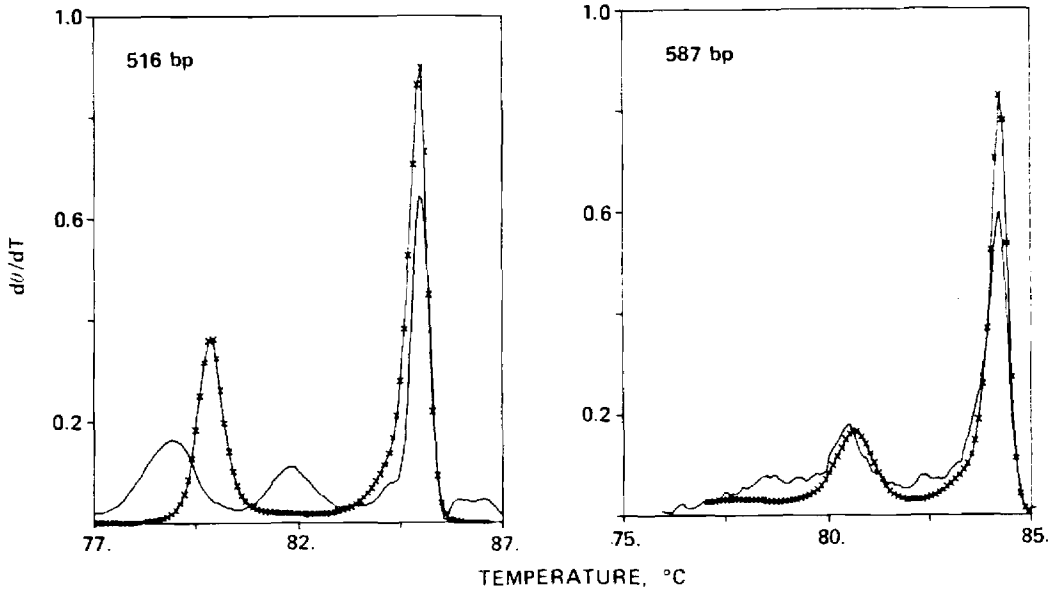


Fig. 18. Theoretical (×) and experimental (——) DMCs of 516 bp and 587 bp DNAs. See table 3 for further information.

AAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TTTGTTTGGTGGCGACCATCGCCACCAAAAAA

TCAAAAAGGATCTTCACCTAGATCCTTTTAAA
AGTTTTTCCTAGAAGTGGATCTAGGAAAATTT

AAAAAAGGATCTCAAGAAGATCCTTTGAT
TTTTTTCCTAGAGTTCTTCTAGGAAACTA

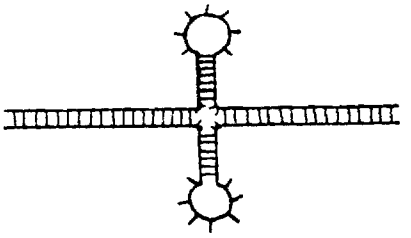


Fig. 19. Three sequences in the 516 bp DNA from pBR322 which can form stem-loop structures.

fragments described in figs. 12 and 13. The shapes of the newly predicted curves were very similar to those depicted in figs. 12 and 13 but the T_m values were consistently higher than previous calculations by 0.5–1.2°C. The experimental melting curves of the *lac* DNAs (except for the 144 bp DNA) were obtained by Wells and collaborators. The pBR322 DNAs and the 144 bp *lac* DNA were obtained by us. Since different instrumentation and analysis procedures were employed, some systematic differences are not surprising.

The studies on the short DNAs are consistent with a set of δG_{MN} values which are small relative to G_S. The observed agreement of 14 out of 15 DNAs, 80–587 bp in length, using two slightly different parameter sets clearly demonstrates the validity of the theory. The lack of agreement noted for the 516 bp DNA can be rationalized in terms of the three stem-loop configurations which can occur in this DNA. The potential for stem-loop structures may well provide the explanation for the difficulty encountered in obtaining consistent theory–experiment agreement for kilo base pair length DNAs. The longer the DNA, the greater the probability of stem-loop structures.

The parameters listed in table 3 are not claimed to be unique, or even an optimized set. It appears unlikely to us, however, that optimization will generate $\delta G_{MN} \geq 200$ for many MN pairs. Predicted melting temperatures and to a lesser extend curve shape are sensitive to changes in $|\delta G_{MN}| \simeq 100$. The experimental data was clearly not compatable with the large δG_{MN}'s of sets 2–4 in table 1. As expected the predicted melting curves were on the whole not sensitive to $\sigma f(m)$. This sensitivity does, however, increase with DNA length. The locations of the two subtransitions in the 587 bp DNA was sensitive to $\bar{\sigma}$ and our analysis indicates that $\bar{\sigma} = 4.5 \times 10^{-5} \pm 50\%$.

Blake and collaborators (personal communication) have initiated a series of studies at determining $f(m)$ by examining the melting of constructed derivatives of pBR322 in 0.08 M Na^+. One DNA has a dA · dT block of about 120 base pairs inserted at the Pvu I restriction site on this DNA. The DNA is linearized at the Eco RI site. This places the A · T block close to the center of the DNA. The first melting subtransition of the plasmid corresponds to the A · T block, and occurs with a T_m which is 5°C higher than for $d(A)_n \cdot d(T)_n$. Analysis of the melting curve suggests that $\sigma \simeq 10^{-6}$ with an exponent of $k = 1.8$ for $f(m) = 1/(m-1)^k$. Further studies of such constructed plasmids will help determine σ and $f(m)$.

3.4. Non-equilibrium effects on short DNA melting

The results described above strongly suggested that the melting curves for short DNAs in 0.1 M Na^+ are close to equilibrium. It is necessary to examine this experimentally. Theoretical analysis by Anshelevich et al. [62] predicts that the strands dissociation step for the short DNAs is irreversible but not highly irreversible. Figure 20 shows a denaturation/renaturation curve for the 144 bp *lac* DNA. The denaturation curve employed a heating rate of 6–7°C/h (no significant difference was noted at 3°C/h). After the DNA was heated about 10°C beyond the transition end point, the temperature was stabilized and then decreased at a rate of 3°C/h. The close agreement between denaturation and renaturation curves indicates the unwinding step is close to reversible. A similar result was obtained with the 457 bp DNA.

It is important to note that the degree of reversibility seen in fig. 20 was not always observed. For example a 157 bp DNA showed marked irreversibility. The latter DNA has 26 and 32 base pairs of repeated (dC–dG) sequences sandwiching the 95 bp *lac* DNA segment [72]. The lack of reversibility may be due to the (dC–dG) regions forming intrastrand stem-loop segments. The above observations indicate that irreversibility in the last step of a melting curve may be due more to stem-loop intermediates created during reannealing than an inherent slow rate of the reassociation step. In the analysis of Anshelevich et al. [62] stem-loop formation is not considered.

A comparison of theory with the experimental melting curve of the 157 bp DNA supports the above notion. Figure 21 shows the experimental melting curve of the 157 bp DNA along with two theoretical curves (Wartell, Benight, Zacharias and Wells, unpublished). The solid line is the experiment. The dashed line is the predicted curve using the parameters described in table 3. The line connecting ×'s is a

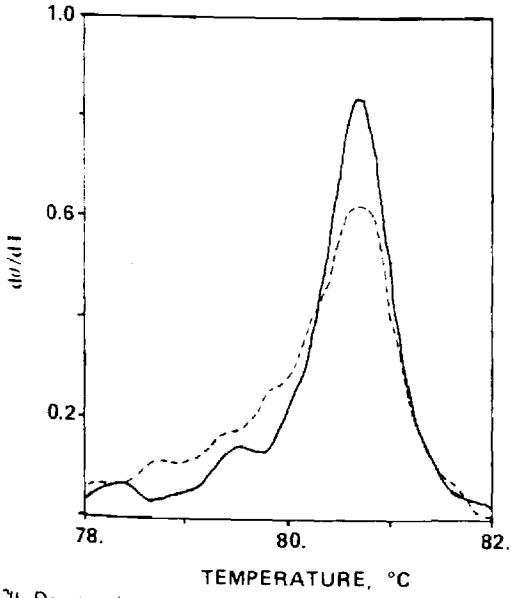


Fig. 20. Denaturation and renaturation differential melting curves are shown for one sample of the 144 bp *lac* DNA. The denaturation curve is the solid line, the renaturation curve the dashed line.

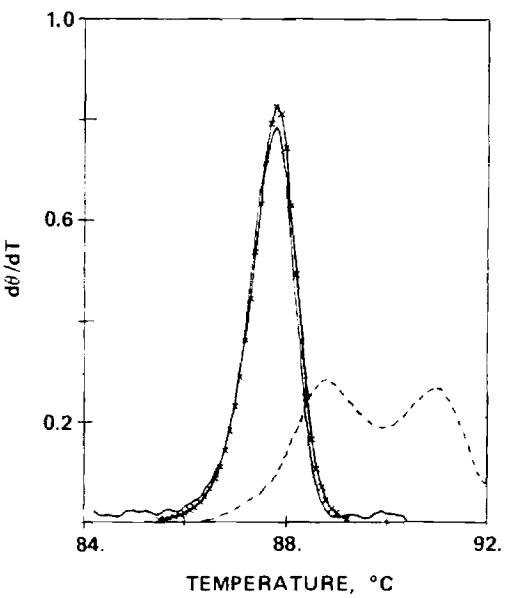


Fig. 21. Experimental DMC of 157 bp DNA in 0.105 M Na^+ is shown as solid line. Dashed line is theoretical curve using parameters given in table 3. Line connecting ×'s is theoretical curve with parameters as in table 3 except $a = -6.3$, $b = 0$. The latter curve was shifted −0.3°C to match experimental curve.

theoretical curve with the same parameters as in table 3 except for the dissociation parameters. These were changed to $b = 0$ and $a = -6.3$.

The effect of these changes is to greatly increase the equilibrium constant for strand separation. One would expect this behavior if kinetically blocked intermediates inhibit strand reassociation. This alteration brings the theoretical curve into excellent agreement with experiment. Decreasing $\tilde{\sigma}$ also converts the bimodal transition curve to a single peak (not shown, but see [53]). However, in this latter case the single peak is over 2°C higher than the experimental T_m. In addition one must decrease $\tilde{\sigma}$ to 10^{-7} in order to achieve a peak height similar to experiment. This value is not compatable with the melting curves of the other short DNAs.

4. The effect of single base pair changes on DNA melting curves

The melting curve studies on the short DNAs indicate that there are small differences in the stacking energies of the ten base pair doublets. A quantitative test of this conclusion may be possible by examining how a DNA fragment melting curve is modified by single base pair changes. Theoretical and experimental results described below indicate that this approach is feasible. The melting curves of two DNAs which differ by one base pair out of several hundred can be distinguished. It is possible to turn this process around and employ DNA melting to screen for small sequence variations in a given DNA restriction fragment. Lerman has demonstrated the potential for this latter concept [7]. He and his colleagues have shown that DNA molecules electrophoresing through a gradient of increasing concentrations of a chemical denaturant, will abruptly decrease in their mobility when they reach a level of denaturant which unwinds part of the DNA. The depth of migration was found to be proportional to the stability of the least stable segment of the DNA. The increasing concentration of chemical denaturant (formamide-urea) appeared to be linearly equivalent to a gradual increase in temperature [73, 74]. Temperature induced DNA melting in solution, and chemically induced melting in a gel are not expected to correspond in detail. However, the similar results of the two procedures suggests that theoretical analysis of how base pair changes effect DNA stability will be helpful in interpreting gel experiments. A method for screening DNA fragments for single base pair changes is of considerable practical importance.

Two studies have been carried out which examined the effect of single base pair changes ('point mutations') or small deletions on DNA melting in solution. Schaeffer et al. [75] compared the melting curves of the 203 bp *lac* DNA fragment containing the normal, or 'wild type' sequence with DNAs containing the UV 5 and/or L8 point mutations. The UV5 mutation changes a G·C to an A·T and an A·T to a T·A. The L8 mutation alters a G·C to A·T. A solvent of 0.01 M Na^+ was employed. Benight and Wartell [53] theoretically examined the effects of 16 known point mutations in three DNA fragments; four mutations in the 203 bp *lac* DNA, eight base pair changes in a 257 bp DNA from the lambda phage P_E promoter, and four mutations in the 192 bp DNA from plasmid pBR322. A solvent of 0.1 M Na^+ was assumed. Fischer and Lerman [73] have employed denaturing gradient gels to show that DNA fragments 536 bp long differing by one base pair can be clearly distinguished.

It is unfortunate that the melting studies of Schaeffer et al. were carried out in 0.01 M Na^+ since this solvent gives irreversible melting behavior. It is of interest nonetheless that theoretical predictions are in good agreement with experiment. Schaeffer et al. and Benight and Wartell predict that the UV5 mutation should decrease the T_m by 0.2–0.5°C and lower the peak of the DMC by about 10%. Experimentally the T_m is lowered by 0.14°C and $(d\theta/dT)_{max}$ is reduced from 1.08 to 0.87. The L8 mutation is predicted to

have a lower T_m and higher value of $(d\theta/dT)_{max}$ from the wild type sequence. Changes observed in the experimental melting curve are in good agreement with these predictions. Theoretical calculations only assumed two base pair stability parameters. Additional calculations by us show that the inclusions of small variations in stacking interactions can result in measurable changes in the DNA melting behavior. The results of Benight and Wartell [53] showed that the effect of a single base pair change varies not only with respect to the type of base pair changed, but also the location of the change and its surrounding sequence. This theoretical conclusion has been demonstrated by Fisher and Lerman [73] in the gel experiments. Sixteen different mutant sequences of a 536 bp DNA fragment were examined. Not only did G·C to A·T or A·T to G·C substitutions result in migration changes, but a G·C to C·G transversion also resulted in a measurably different gel mobility. A linear correlation was observed between the depth of migration of seven fragments relative to the wild type fragment and the T_m calculated for the lowest melting region. The Fixman–Friere method was used with the Gotoh and Tagashira parameters. The authors acknowledge that other nearest neighbor stacking parameters may also be compatible with this correlation.

5. Fluctuational base pair opening at temperatures below the transition

The previous sections of this review have emphasized experimental results on DNA helix-coil transitions and the development of a thermodynamic model which describes the transitions. Biological processes involving DNA generally occur at 25–37°C, far below the transition region. At these lower temperatures, DNA is predominantly a hydrogen bonded double helix. Consideration of Brownian motion of solvent and DNA dictate that spontaneous fluctuational opening of base pairs will occur. Although the probability of base pair opening is expected to be considerably smaller than in the transition region, it may be highly significant with regard to protein–DNA binding. RNA polymerase, for example, can bind to DNA in 0.1 M NaCl at 37°C and induce a region of 10–15 bp to unwind. An understanding of this process will require an understanding of DNA base pair opening below the transition. Does the RNA polymerase loosely bind to the DNA and await the ambient fluctuational opening of one or more base pairs, or does its binding interaction affect a lowering of the DNA base pair stability? What is the probability that a base pair will open in 0.1 M Na^+ at 25 or 37°C? What is the influence of a base pair's sequence environment on its probability of opening?

A number of workers have been concerned with these and related questions. Different experimental measurements of 'base pair opening' have generated a range of values. This is due to the difficulties and uncertainties inherent in the approaches employed, and also because different methods are probing different open states. The open base pair state we are considering may be viewed as an average over conformations in which base pair hydrogen bonds are broken, and the two bases do not have the complete stacking interactions which occur in the duplex state. This, of course, is a vague conceptual picture. Operationally we define an open base pair as a conformation which results in a fractional increase in UV absorbance near or at 268 nm in a DNA melting curve. This definition allows us to employ results obtained from UV denaturation data to examine base pair opening. We assume that this 'UV-absorbance' open state allows a chemical group to probe base atoms normally involved in Watson–Crick hydrogen bonding.

Base pair opening as measured by the exchange of hydrogen bonded imino protons on the bases with deuterium or tritium [76, 77] probe a different type of open state than the one we consider. Environ-

mental variables differentially effect the hydrogen exchange open state and UV absorbance open state [78]. NMR measurements of proton spin relaxation by chemical exchange give values similar to the hydrogen–deuterium exchange data. The nature of this hydrogen exchange open state is also uncertain. Results indicate it does not involve any base unstacking and this open state will not be considered further.

The theoretical formalism employed to describe base pair opening in the helix-coil region can be used to describe base pair opening at low temperatures. The probability that the ith base pair in a specific DNA sequence is open, $\theta_1(i)$ is obtained directly from the Poland calculational method. One has from eq. (25)

$$\theta_1(i) = 1 - p(1) . \tag{43}$$

It is of interest to characterize the behavior of θ_1 as a function of temperature. This was first done by Lukashin et al. [78]. In the region of the helix-coil transition, $T \simeq T_m$, base pairs melt as large cooperative regions (e.g., fig. 6). Thus the value of $\theta_1(i)$ is strongly influenced by the overall stability of the ith base pair's surrounding sequence. For temperatures below the transition, this correlated behavior gradually decreases.

Figure 22 shows plots of $\theta_1(i)$ vs. i for a 60 bp region of the 203 bp *lac* DNA 10°C and 48°C below the T_m. The region depicted is the promotor region or RNA polymerase transcription binding site for the lactose operon of E. coli. The quantitative values of θ_1 must be viewed with caution (see below). We are here only concerned with the qualitative behavior of θ_1 as a function of temperature. At 10°C below the T_m θ_1 for A · T base pairs varies by 220%. Base pair opening depends on the nearby base pair sequences. This range decreases to a 6% variation at 48°C below the T_m. At this temperature base pair opening depends mainly on the base pair type (A · T or G · C) and not its neighbors. It is clear from fig. 22 that the length over which a base pair influences the opening of other base pairs is temperature dependent. The dotted lines in fig. 22 show the alterations in θ_1 when G · C pairs at positions −16 or −9 are individually changed to A · T pairs. The effect at 10°C below the T_m is much larger than at 48°C below the T_m.

The above predictions can be expected to be qualitatively correct, but may be in considerable error quantitatively. This point was discussed in detail by Wartell and Benight [24] and is briefly reviewed. The calculation of θ_1 will be accurate at low temperatures only if the nearest neighbor model accurately describes the actual physical interactions along DNA. Poland and Scheraga [11] pointed out that the nearest neighbor formalism cannot be distinguished from a model with longer range interactions when the melting of large regions of base pairs dominates the partition function. This, of course, is the situation in the helix-coil region. If longer range interactions occur the stacking parameter σ actually averages in any long range interactions which can occur at helix-coil junctions. As the size of the average loop of unpaired bases decreases, σ may become sensitive to interactions beyond nearest neighbor which are negligible in the transition region. The nearest neighbor formalism can be preserved, but σ will appear to change with loop size. One example is the stacking interaction which may occur between base pairs i and $i + 2$ when base pair $i + 1$ opens. Patel et al. [79] observed this type of interaction when one strand has an extra unpaired base. When large loops dominate the partition function this interaction can be neglected. At low temperatures, this next neighbor interaction may be important. In addition to the above effect, $f(m)$ for small m can be expected to be sequence dependent (see section 2.1.3).

Several experiments have been directed at evaluating the probability of single base pair opening

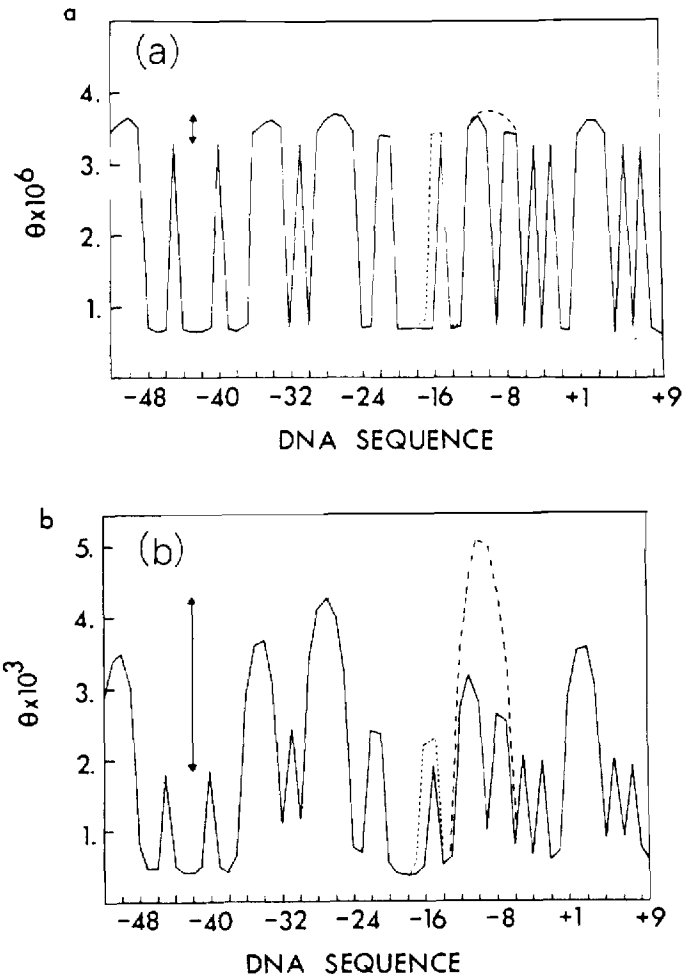


Fig. 22. The probability of base pair opening θ_1 is plotted for the *lac* promoter region. (a) Profile at 48°C below the T_m of the DNA, (b) 10°C below the T_m of the DNA. The range of θ_1 for A·T pairs is shown by ↔.

[78, 80, 81]. Gralla and Crothers evaluated $\theta_1(i)$ from the melting of RNA duplexes in 1.0 M NaCl [80]. RNA oligomers with completely complementary base pairs were compared with similar oligomers containing opposing unpaired cytosines. Assuming the cytosines were unstacked and formed an interior loop, they estimated the free energy required to form a one base pair loop. Lukashin et al. [78] and McGhee and von Hippel [81] employed formaldehyde induced denaturation of DNA to evaluate fluctuational base pair opening in 0.02–0.2 M NaCl. The experiments of Lukashin et al. were in agreement with the predictions based on helix-coil transition parameters. However, as discussed elsewhere, these experiments monitored the extension of previously formed loops [24, 81]. Lukashin et al. acknowledge that their agreement "cannot guarantee the correctness of calculated values . . . for very small loops." The results obtained by McGhee and von Hippel, and Gralla and Crothers if extrapolated with certain assumptions, indicate a probability of base pair opening 10^2 to 10^3 greater than predicted by the helix coil parameters [24]. This apparent paradox may indicate the inadequacy of the nearest neighbor model as discussed earlier. There are however a number of uncertainties in the evaluation of $\theta_1(i)$ from these experiments which make the determined values less than unequivocal.

The Gralla Crothers work evaluated $\theta_1(i)$ for RNA in 1.0 M NaCl. A number of assumptions must be made in order to extrapolate the data to DNA base pair opening in 0.1 M or lower Na^+. The most important assumption concerns the use of a mismatched pair of bases as a model for the open base pair state. Tibanyenda et al. [33] recently examined the effect of six separate base pair mismatches on the melting curve of a 16 bp DNA duplex. T–C, A–C, T–G, and A–A mismatches were examined. NMR and thermodynamic analysis indicated that stacking interactions occur between the two mispaired bases and neighboring base pairs. NMR studies by Patel et al. [79] also indicate that a mismatched base may or may not stack within adjacent base pairs depending on the local base pair sequence. Mispaired bases are not in general accurate models for an open unstacked base pair. Other concerns regarding the application of the RNA data, such as differences in stability between RNA and DNA, and the corrections to the free energies required to extrapolate from 1.0 M NaCl to 0.1 M NaCl create further uncertainty. The agreement of the Gralla Crothers data with the results of McGhee and von Hippel may be fortuitous.

The work of McGhee and von Hippel is more directly involved with DNA base pair opening, and is the best available experimental estimate. An estimate of $\theta_1(i)$ was obtained from the reaction of formaldehyde with the exocyclic nitrogen of adenine, N6, in poly (dA–dT) · poly (dA–dT) at temperatures 20°C or more below the T_m. A detailed investigation of the competing reaction of formaldehyde with the endocyclic nitrogen N3 of thymine, and the rate of formaldehyde's reaction with single stranded polynucleotides, together with other considerations led these authors to conclude that they were measuring the reaction of formaldehyde with an open unstacked adenine. Studies on natural DNAs were consistent with their findings on poly (dA–dT) · poly (dA–dT). Several aspects of the analysis has caused Frank–Kamenetskii [82] and Wilcoxson and Schurr [83] to raise doubts as to its validity. Although a number of criticisms can be readily deflated [84], the formaldehyde DNA reaction is a complex one, and interpretation of the data is not straightforward. The large difference between helix-coil theory and the experimental value for base pair opening, and the possibility that formaldehyde may be reacting with a stacked adenine at low temperatures [83] suggests the need for further investigations. New experimental approaches are required to quantify fluctuational base pair opening.

Acknowledgements

The authors wish to acknowledge the excellent typing of Audrey Ralston. This work was supported in part by a grant from the National Institutes of Health which is gratefully acknowledged.

Appendix

Absorbance vs. temperature measurements were made by monitoring UV absorbance of DNA samples at 268 nm while slowly increasing the temperature. A Beckman double-beam Acta MVI spectrophotometer with digital output capability was employed. Variation in the absorbance was ±0.0005 over a 10 min interval. The DNA sample (1 ml, 0.4–0.5 OD/ml at 268 nm) and an equal volume of solvent were placed in identical cuvettes and inserted in symmetrically adjacent positions in the cell holder. The cell holder was heated by circulating through its channels a 30% ethylene glycol solution from a Lauda temperature controlled circulator. Sample temperatures were measured by a platinum resistance thermometer placed and sealed in the solvent cuvette. The standard deviation of the

temperature of the major peak of the differential melting curves was about ±0.3°C for three experiments of a given DNA.

Prior to the melting curve studies, DNA samples were dialyzed against 0.1 M NaCl + 1 mM potassium phosphate (pH 7.4) + 10 mM Na_2 EDTA. Helium gas was bubbled through the samples to displace dissolved air. The temperature of a sample was raised at about 7°C/h for most experiments. A rate of about 3°C/h did not result in any significant differences. Absorbance, temperature readings were made at approximately 0.05°C intervals. Data was collected on paper tape, and then transferred to and processed by the Georgia Tech CDC time sharing computer.

The analysis procedure was similar to that described by Ansevin et al. [26]. The absorbance data were first corrected for volume expansion of the solution and then interpolated to equally spaced temperature intervals. Linear equations for $A_L(T)$ and $A_U(T)$ were determined graphically and an initial $\theta_B(T)$ calculated (see eq. (1) and fig. 2). The Savitsky–Golay method was used to smooth the data [86]. A cubic polynomial was first fit to $\theta_B(T)$ using 13 point intervals. In this process a least squares fit is made over a 13 point interval and the midpoint of the cubic curve replaces the original θ_B value at this temperature. The 13 point window is moved one data point and the process continued. The derivative of the once smoothed curve was then subjected to a second round of smoothing. A cubic polynomial was fit over a 9 point interval. The number of points selected for the smoothing intervals influences the final value of $(d\theta/dT)_{max}$. Earlier analysis [87] utilizing 15 and 21 point windows resulted in lower values of $(d\theta/dT)_{max}$ and slightly broader peaks. The intervals chosen in this work appeared to give the minimum distortion of peak heights and shapes when compared to unsmoothed data. The smoothed differential melting curves of several experiments (2–3) were aligned in temperature and then averaged. Reproducibility of $(d\theta/dT)_{max}$ was ±10% and full width at half height of transition peaks, $\Delta T_{1/2} = \pm 0.01$°C. The internal difference in T_m values for transitions which showed two clear peaks was ±0.03°C.

References

[1] R. Thomas, Biochim. Biophys. Acta 14 (1954) 231.
[2] S.A. Rice and P. Doty, J. Amer. Chem. Soc. 79 (1957) 3937.
[3] W. Szybalski, Thermobiology, A.H. Rose, ed. (Acad. Press, London, 1967).
[4] D.M. Crothers, Accts. Chem. Res. 2 (1969) 225.
[5] G. Newell and E.W. Montroll, Rev. Mod. Phys. 25 (1953) 353.
[6] B.H. Zimm, J. Chem. Phys. 33 (1960) 1349.
[7] L.S. Lerman, S.G. Fischer, I. Hurley, K. Silverstein and N. Lumelsky, Ann. Rev. Biophys. 13 (1984) 399.
[8] N. Nishigaki, Y. Husimi, M. Masuda, K. Kaneko and T. Tanaka, J. Biochem. 95 (1984) 627.
[9] A. Wada, S. Yubuki and Y. Husimi, CRC Crit. Rev. Biochem. 9 (1980) 87.
[10] O. Gotoh, Adv. Biophys. 16 (1983) 1.
[11] D. Poland and H.R. Scheraga, Theory of Helix Coil Transition in Biopolymers (Acad. Press, New York, 1970).
[12] Y.S. Lazurkin, M.D. Frank-Kamenetskii and E.N. Trifonov, Biopolymers 9 (1970) 1253.
[13] R.M. Wartell and E.W. Montroll, Adv. Chem. Phys. 22 (1972) 129.
[14] V.A. Bloomfield, D.M. Crothers and I. Tinoco, Physical Chemistry of Nuc. Acids (Harper and Row, New York, 1974).
[15] S. Hirschman and G. Felsenfeld, J. Mol. Biol. 6 (1966) 347.
[16] J. Marmur and P. Doty, J. Mol. Biol. 5 (1960) 109.
[17] R.J. Owen, L.R. Hill and S.P. LaPage, Biopolymers 7 (1969) 503.
[18] S. Falkow and D. Cowie, J. Bacterial. 96 (1968) 777.
[19] S. Yabuki, M. Fuke and A. Wada, J. Biochem. 69 (1971) 191.
[20] Y. Lyubchenko, A.V. Volgodskii and M.D. Frank-Kamenetskii, Nature 271 (1978) 28.
[21] V.M. Pavlov, J.L. Lyubchenko, A.S. Borovik and Y. Lazurkin, Nucl. Acids. Res. 4 (1977) 4053.
[22] A.S. Borovik, Y.A. Kalambet, Y.L. Lyubchenko, V.T. Shitov and E. Golovanov, Nuc. Acids. Res. 8 (1980) 4165.

[23] A.S. Benight, D.K. Howell and R.M. Wartell, Nature 289 (1981) 203.
[24] R.M. Wartell and A.S. Benight, Biopolymers 21 (1982) 2069.
[25] V.I. Lyanichev, I. Panyutin, D. Cherny and Y.L. Lyubchenko, Nuc. Acids Res. 11 (1983) 2166.
[26] D. Vizard, R. White and A.T. Ansevin, Nature 275 (1978) 250.
[27] O. Gotoh and Y. Tagashira, Biopolymers 20 (1981) 1033.
[28] P.N. Borer, B. Dengler, I. Tinoco and O.C. Uhlenbeck, J. Mol. Biol. 86 (1974) 843.
[29] L.A. Marky and K.J. Breslauer, Biopolymers 21 (1982) 2185.
[30] R.D. Wells and R.M. Wartell, Biochemistry of Nuc. Acids 6 (1974), H.L. Kornberg, D.C. Philips and K. Burton, eds. (Univ. Park Press, Balt., 41).
[31] R.L. Ornstein and R. Rein, Biopolymers 18 (1979) 2821.
[32] R.L. Ornstein and J.R. Fresco, Biopolymers 22 (1983) 1979.
[33] N. Tibanyenda, S. DeBruin, C. Haasnoot, G. van der Marel, J. Van Boom and C. Hilbers, Eur. J. Biochem. 139 (1984) 19.
[34] H. Klump and T. Ackermann, Biopolymers 10 (1971) 513.
[35] B.N. Belenktsev, A.V. Volgodskii and M.D. Frank-Kamenetskii, Mol. Biol. 10 (1976) 764.
[36] J. Josse, A.D. Kaiser and A. Kornberg, J. Biol. Chem. 236 (1961).
[37] D.W. Marquardt, J. Soc. Ind. Appl. Math. 11 (1963) 431.
[38] T.R. Krugh and C.G. Reinhardt, J. Mol. Biol. 97 (1975) 133.
[39] H. Jacobson and W.H. Stockmayer, J. Chem. Phys. 18 (1950) 1600.
[40] M. Fixman and J. Friere, Biopolymers 16 (1977) 2693.
[41] (a) M.Y. Azbel, Biopolymers 19 (1980) 61;
(b) M.Y. Azbel, Biopolymers 19 (1980) 81.
[42] C.R. Cantor and P. Schimmel, Biophysical Chemistry, Vol. III (Freeman, San Francisco, 1980).
[43] R.M. Wartell, Nuc. Acids Res. 4 (1977) 2779.
[44] G. Felsenfeld and H.T. Miles, Ann. Rev. Biochem. 36 (1967) 407.
[45] D. Poland, Biopolymers 13 (1974) 1859.
[46] R.H. Lacombe and R. Simha, J. Chem. Phys. 58 (1973) 1043.
[47] M.D. Frank-Kamenetskii and A.D. Frank-Kamenetskii, J. Mol. Biol. 3 (1969) 295.
[48] Y.L. Lyubchenko, A. Kalambet, V.I. Lyamichev and A.S. Borovik, Nuc. Acids Res. 10 (1982) 1867.
[49] M. Wadati and A. Ishihara, Mol. Cryst. Liq. Cryst. 17 (1972) 95.
[50] E.W. Mayer and M.G. Mayer, Statistical Physics, 2nd ed. (Wiley, New York, 1977) p. 329.
[51] L.D. Landau and E.M. Lifschitz, Statistical Physics (Addison Wesley, Mass., 1958) p. 148.
[52] P. Hagarman and B.H. Zimm, Biopolymers 20 (1981) 1481.
[53] A.S. Benight and R.M. Wartell, Biopolymers 22 (1983) 1409.
[54] B.Y. Tong and S.J. Battersby, Biopolymers 18 (1979) 1917.
[55] N. Panayotatos and R.D. Wells, Nature 289 (1981) 466.
[56] D.M.J. Lilley, Proc. Natl. Acad. Sci. 77 (1980) 6468.
[57] S. Yabuki, M. Fuke and A. Wada, J. Biochem. 89 (1971) 191.
[58] F. Michel, J. Mol. Biol. 89 (1974) 305.
[59] M.P. Perelroyzen, V.I. Lyamichev, A.Y. Kalambet, L.Y. Lyubchenko and A.D. Volgodskii, Nuc. Acids Res. 9 (1981) 4043.
[60] P.V. Haydock, Ph.D. Thesis, University of Maine, Orono, Ma. (1979).
[61] H. Tachibana, S. Ueno-Nishio, O. Gotoh and A. Wada, J. Biochem. 92 (1982) 623.
[62] V.V. Anshelevich, A.V. Vologodskii, A.V. Lukashin and M.D. Frank-Kamenetskii, Biopolymers 23 (1984) 39.
[63] (a) M.E. Craig, D.M. Crothers and P. Doty, J. Mol. Biol. 62 (1971) 383.
(b) D. Porschke and M. Eigen, J. Mol. Biol. 62 (1971) 361.
[64] J. Gabarro-Arpa and F. Michel, Biochimie 64 (1982) 99.
[65] H.C. Spatz and D.M. Crothers, J. Mol. Biol. 42 (1969) 191.
[66] A.J. Hoff and A.L.M. Roos, Biopolymers 11 (1972) 1289.
[67] A. Suyama and A. Wada, Biopolymers 23 (1984) 409.
[68] R.D. Wells, T.C. Goodman, W. Hillen, G.T. Horn, R.D. Klein, J. Larson, U. Muller, S.K. Neuendorf, N. Panayotatos and S.M. Stirdivant, Prog. Nuc. Acid Res. 24 (1980) 167.
[69] S.C. Hardies and R.D. Wells, Gene 7 (1979) 1.
[70] S.C. Hardies, W. Hillen, T.C. Goodman and R.D. Wells, J. Biol. Chem. 254 (1979) 10128.
[71] W. Hillen, T. Goodman, A.S. Benight, R.M. Wartell and R.D. Wells, J. Biol. Chem. 256 (1981) 2761.
[72] J. Klysik, S. Stirdivant, J. Larsen, P.A. Hart and R.D. Wells, Nature 290 (1981) 671.
[73] L. S. Lerman, S. G. Fischer, D.B. Bregman and K. Silverstein, Biomole. Stereodynamics, R.H. Sarma, ed. (Adenine, New York, 1981) p. 459.
[74] S.G. Fischer and L.S. Lerman, Proc. Natl. Acad. (USA) 80 (1983) 1579.
[75] F. Schaeffer, A. Kolb and H. Buc, EMBO J. 1 (1982) 99.
[76] H. Tenelbaum and S.W. Englander, J. Mol. Biol. 92 (1975) 55.

[77] N.R. Kallenback, C. Mandal and S.W. Englander, Nuc. Acid Geometry and Dynamics, R.H. Sarma, ed. (Pergamon Press, New York, 1980) p. 233.
[78] A.V. Lukashin, A.V. Volgodskii, M.D. Frank-Kamenetskii and Y. Lyubchenko, J. Mol. Biol. 108 (1976) 665.
[79] D. Patel, S.A. Kozlowski and D.R. Hare, Biochemistry 23 (1984) 3207.
[80] J. Gralla and D.M. Crothers, J. Mol. Biol. 78 (1973) 301.
[81] J.D. McGhee and P.H. von Hippel, Biochemistry 16 (1977) 3276.
[82] M.D. Frank-Kamenetskii, Comm. Mol. and Cell. Biophysics 1 (1981) 105.
[83] J. Wilcoxson and M. Schurr, Biopolymers 22 (1983) 2273.
[84] J.D. McGhee and P.H. von Hippel, personal communication.
[85] J.G. Sutcliffe, Cld. Spg. Harbor Symp. Quant. Biol. 43 (1978) 77.
[86] A. Savitsky and M.J.E. Golay, Anal. Chem. 36 (1964) 1627.
[87] A.S. Benight, Ph.D. Thesis, Georgia Institute of Technology, 1983.